

IT 00 / 48



09/913878

MINISTERO DELL'INDUSTRIA, DEL COMMERCIO E DELL'ARTIGIANATO**DIREZIONE GENERALE DELLA PRODUZIONE INDUSTRIALE
UFFICIO ITALIANO BREVETTI E MARCHI**

REC'D 07 JUN 2000

WIPO

PCT

Autenticazione di copia di documenti relativi alla domanda di brevetto per INV. IND.**N. RM99A000117 DEL 22.02.1999**

*Si dichiara che l'unita copia è conforme ai documenti originali
depositati con la domanda di brevetto sopraspecificata, i cui dati
risultano dall'accluso processo verbale di deposito*

**PRIORITY
DOCUMENT**
SUBMITTED OR TRANSMITTED IN
COMPLIANCE WITH RULE 17.1(a) OR (b)

02 MAG 2000

Roma, li

IL DIRETTORE DELLA DIVISIONE

Ing. DI CARLO

A. RICHIEDENTE (I)

1) Denominazione UNIVERSITA' DEGLI STUDI DI ROMA "LA SAPIENZA" N.G. EN
Residenza Roma, RM codice 02133771002
2) Denominazione _____
Residenza _____ codice _____

B. RAPPRESENTANTE DEL RICHIEDENTE PRESSO L'U.I.B.M.

cognome nome CAPASSO Olga ed altri cod. fiscale _____
denominazione studio di appartenenza ING. BARZANO' & ZANARDO ROMA S.p.A.
via Piemonte n. 26 città Roma cap 00187 (prov) RM

C. DOMICILIO ELETTIVO destinatario

via _____ n. _____ città _____ cap _____ (prov) _____

D. TITOLO

classe proposta (sez/cl/sci) _____

gruppo/sottogruppo _____

"Isolamento e caratterizzazione di un gene per il silenziamento genico da N. crassa e suoi usi".

ANTICIPATA ACCESSIBILITÀ AL PUBBLICO: SI ☐ NO ☒

SE ISTANZA: DATA _____ N° PROTOCOLLO _____

E. INVENTORI DESIGNATI

cognome nome

cognome nome

1) COGONI Carlo 3) _____
2) MACINO Giuseppe 4) _____

F. PRIORITÀ

nazione o organizzazione

tipo di priorità

numero di domanda

data di deposito

allegato S/R

1) _____
2) _____

SCIoglimento RISERVE

Data

N° Protocollo

G. CENTRO ABILITATO DI RACCOLTA COLTURE DI MICRORGANISMI, denominazione

H. ANNOTAZIONI SPECIALI

Nessuna

DOCUMENTAZIONE ALLEGATA

N. es.

Doc. 1) 12 PROV n. pag. 41 riassunto con disegno principale, descrizione e rivendicazioni (obbligatorio 1 esemplare)
Doc. 2) 2 PROV n. tav. 05 disegno (obbligatorio se citato in descrizione, 1 esemplare
Doc. 3) 1 RIS lettera d'incarico, ~~lettera di incarico per la preparazione del brevetto~~
Doc. 4) 1 RIS designazione inventore
Doc. 5) 1 RIS documenti di priorità con traduzione in italiano
Doc. 6) 1 RIS autorizzazione o atto di cessione
Doc. 7) 1 nominativo completo del richiedente

SCIoglimento RISERVE

Data

N° Protocollo

confronta singole priorità

8) attestati di versamento, totale lire Cinquecentosessantacinquemila.- obbligatorio

COMPILATO IL 22/02/1999 FIRMA DEL(I) RICHIEDENTE (I) UNIVERSITA' DEGLI STUDI DI ROMA "LA SAPIENZA"

CONTINUA SI/NO NO ING. BARZANO' & ZANARDO ROMA S.p.A. UN MANDATARIO

DEL PRESENTE ATTO SI RICHIEDE COPIA AUTENTICA SI/NO SI Olga Capasso per se e per gli altri

Olga Capasso
(N° d'istr. 820 B)

UFFICIO PROVINCIALE IND. COMM. ART. DI RM 99 A 000117 Roma codice 58

VERBALE DI DEPOSITO NUMERO DI DOMANDA _____ Reg. A

L'anno millenovecento Novantanove, il giorno Ventidue, del mese di Febbraio

il(i) richiedente(i) sopraindicato(i) ha(hanno) presentato a me sottoscritto la presente domanda, costituita di n. 00 fogli aggiuntivi per la concessione del brevetto soprariportato.

I. ANNOTAZIONI VARIE DELL'UFFICIALE ROGANTE

IL DEPOSITANTE

L'UFFICIALE ROGANTE

L'Ufficiale Rogante
Silvia Altieri

RO BREVETTO

RM 99 A 00117

DATA DI RILASCIO

CHIEDENTE(I)

denominazione

UNIVARSITA' DEGLI STUDI DI ROMA "LA SAPIENZA"

denominazione

TOLO

amento e caratterizzazione di un gene per il silenziamento genico da N. crassa e suoi usi".

proposta (sez./cl./scl/)

(gruppo/sottogruppo)

ASSUNTO

E' descritta una sequenza nucleotidica codificante per una proteina caratterizzata dal fatto di avere attività di silenziamento e di comprendere un dominio di RNA polimerasi RNA-dipendente; sono descritti vettori di espressione in grado di esprimere la sequenza in batteri, piante animali e funghi, nonché organismi trasformati con tali vettori.

DISEGNO



DESCRIZIONE

a corredo di una domanda di brevetto per invenzione industriale dal titolo: "Isolamento e caratterizzazione di un gene per il silenziamento genico da *N. crassa* e suoi usi"

a nome: Università degli Studi di Roma "La Sapienza"

inventori: Cogoni Carlo, Macino Giuseppe

La presente invenzione concerne l'isolamento e la caratterizzazione di un gene per il silenziamento genico da *N. crassa* e suoi usi.

Più in particolare l'invenzione riguarda l'isolamento e la caratterizzazione di un gene da *Neurospora crassa* codificante per un'attività essenziale per il processo di co-soppressione, e suoi usi e applicazioni in campo vegetale, animale e fungino.

L'ottenimento di organismi transgenici è di grande utilità sia nella ricerca biologica di base che in quella applicata. Il DNA transgenico viene normalmente integrato nel genoma e trasmesso come carattere Mendeliano. Tuttavia, in diversi casi, l'introduzione di transgeni induce fenomeni di silenziamento genico (Flavell, R.B. 1994), cioè di repressione dell'espressione del transgene stesso e/o di uno o più geni omologhi endogeni.

Il silenziamento genico agisce a due livelli: trascrizionale (trans-inactivation) quando i transgeni contengono sequenze omologhe al promotore del gene silenziato (Vaucheret, 1993); e post-

trascrizionale (co-soppressione) che richiede omologie tra le regioni codificanti (Flavell, 1994; Stam et al. 1997; Baulcombe, 1996).

In generale, il silenziamento indotto da transgeni necessita di una quasi completa omologia di sequenza (tra 70% e 100%) tra il transgene e il gene silenziato (Elkind, 1990).

Nel fungo filamentoso *Neurospora crassa*, in fase vegetativa, la presenza di transgeni induce un fenomeno di silenziamento genico a livello post-trascrizionale, denominato "quelling" (Cogoni et al. 1996).

Attraverso l'uso del gene al-1 (albino 1) (Schmidhauser et al. 1990) come marcatore visuale per il silenziamento, ne sono state evidenziate diverse caratteristiche (Cogoni et al. 1996). In particolare, il "quelling" del gene al-1 in *Neurospora* è caratterizzato da: 1) il silenziamento genico è reversibile, a causa della perdita di copie dei transgeni; 2) la riduzione del livello basale del mRNA è dovuto a un effetto post-trascrizionale; 3) i transgeni contenenti almeno una regione di 132 coppie di basi identica alla regione codificante del gene bersaglio sono sufficienti ad indurre il "quelling"; 4) la duplicazione delle sequenze promotore è inefficace nell'indurre il silenziamento; 5) il "quelling" ha un comportamento dominante in eterocarion contenenti sia nuclei transgenici che nuclei non trasformati, indicando il coinvolgimento di una molecola diffusibile, che agisce "in trans" tra i nuclei; 6) l'espressione di un RNA aberrante trascritto dal locus transgenico è strettamente correlata con il silenziamento, suggerendo che il "quelling" può essere indotto e/o mediato da una molecola di RNA transgenica.

Pertanto possono essere evidenziate delle omologie tra il silenziamento in *Neurospora* e la co-soppressione nelle piante. Il silenziamento genico é reversibile in *Neurospora*, come conseguenza della instabilità delle copie transgeniche in fase mitotica; anche nelle piante la reversione della co-soppressione é associata con la riduzione del numero di copie di transgene, risultante dalla ricombinazione intracromosomale durante la mitosi o la meiosi (Mittelstein Scheid et al. 1994; Stam et al. 1997). Quindi, sia nelle piante che in *Neurospora*, la persistenza dei transgeni é necessaria per il mantenimento del silenziamento. Come in *Neurospora*, la diminuzione del livello basale del mRNA del gene silenziato é causata da un meccanismo post-trascrizionale (Dehio and Schell 1994; van Blokland et al. 1994; de Carvalho et al. 1995). Ed ancora, per l'induzione del "quelling" é necessario che i transgeni contengano una porzione della regione codificante del gene bersaglio del silenziamento, essendo la regione promotore inefficace per la sua induzione. Nelle piante porzioni codificanti senza promotore possono indurre il silenziamento (van Blokland et al. 1994) e, come, nel "quelling" i promotori o i prodotti genici funzionanti non sono richiesti per la co-soppressione.

Uno dei paralleli tra il "quelling" e la co-soppressione nelle piante é che entrambi i meccanismi sono mediati da fattori diffusibili. In ceppi eterocariotici di *Neurospora*, i nuclei in cui il gene albino 1 é silenziato sono in grado di indurre il silenziamento del gene al-1 residente in nuclei non trasformati, ma che condividono un citoplasma comune (Cogoni et al. 1996). Nelle piante la presenza di un fattore

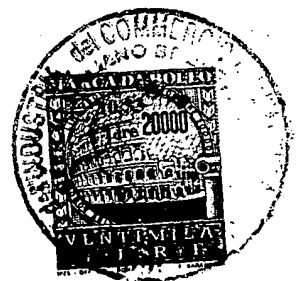
diffusibile é dimostrata dal fatto che la co-soppressione é in grado di inibire la replicazione del Tobacco Etch Virus (TEV), un virus ad RNA con un ciclo esclusivamente citoplasmatico. L'esistenza di fattori altamente diffusibili in grado di mediare la co-soppressione é stata dimostrata usando la tecnica di innesto in tabacco (Palaqui et al., 1997), dimostrando che piante di tabacco silenziate sono in grado di trasmettere il silenziamento a piante non silenziate attraverso un innesto anche a grande distanza.

Il fatto che il "quelling" e la co-soppressione condividano tutte queste caratteristiche suggerisce che i meccanismi coinvolti nel silenziamento genico post-trascrizionale nelle piante e nei funghi può essere evoluto da un comune meccanismo ancestrale.

Recentemente sono stati descritti fenomeni di inattivazione genica in seguito ad introduzione di transgeni anche in specie animali. In *Drosophila melanogaster*, il posizionamento di un transgene vicino a centri eterocromatinici determina una espressione variegata (Wallrath and Elgin 1995; Pirrotta, V. 1997). Simili profili d'espressione sono stati osservati quando il transgene di riferimento é all'interno di trasposoni organizzati in tandem, indicando che le ripetizioni sono in grado di indurre la condensazione della cromatina (Dorer and Henikoff, 1994). Sempre in *Drosophila*, Pal-Bhadra et al. (1997) hanno osservato che l'introduzione di transgeni può condurre a fenomeni di inattivazione genica, simili alla co-soppressione.

Fenomeni di silenziamento genico dovuto a sequenze transgeniche ripetitive sono stati recentemente riportati nei mammiferi.

ING. GABRILO ZANUCCI T.M.A. SpA



Garrick et al. (1998) hanno generato delle linee transgeniche di topo in cui 100 copie transgeniche sono presenti in un unico locus ed organizzate in tandem diretti. E' stato osservato che l'espressione del transgene é inversamente proporzionale al numero di copie presenti, indicando che fenomeni di silenziamento dipendenti da copie ripetute sono presenti anche nei mammiferi.

L'identificazione di geni coinvolti nel silenziamento in *Neurospora* è pertanto il primo passaggio per modulare lo stesso processo in piante, animali e funghi. La modulazione del silenziamento è di grande importanza applicativa per la produzione di organismi transgenici in grado di esprimere il fenotipo di interesse.

Gli autori della presente invenzione hanno già isolato ceppi di *Neurospora crassa* mutati in funzioni essenziali per il meccanismo di silenziamento genico (Cogoni and Macino, 1997): 15 mutanti indipendenti isolati definiscono tre gruppi di complementazione identificando tre geni *qde-1*, *qde-2* e *qde-3* (per *qde* si intenda "quelling" deficient), i cui prodotti sono indispensabili per il macchinario di silenziamento. I geni *qde* sono essenziali per il silenziamento in *Neurospora*, come suggerito dal fatto che il silenziamento di tre geni indipendenti (*al-1*, *al-2* e *qa-2*) é impedito dalle mutazioni *qde* (Cogoni and Macino, 1997).

Gli autori della presente invenzione hanno già identificato il gene *qde-3* (Domanda di Brevetto Italiano No. RM98/A000730).

Gli autori della presente invenzione hanno ora identificato e clonato un altro dei geni *qde* in *Neurospora*, il gene *qde-1*,

identificando un altro dei fattori richiesti per il silenziamento. In considerazione delle similarità tra "quelling" e co-soppressione, i geni ortologhi al gene isolato sono coinvolti nella co-soppressione e più in generale nel silenziamento genico in altri organismi, sia piante, che funghi, che animali.

Si possono individuare due ambiti generali per l'applicazione dell'invenzione: 1) potenziamento del silenziamento come sistema per inattivare in maniera più efficace e stabile un gene d'interesse, e 2) soppressione del silenziamento per ottenere una migliore espressione dei transgeni introdotti.

Per un potenziamento del silenziamento, la sovraespressione di uno o più geni che controllano il fenomeno può condurre ad una maggiore efficienza e/o stabilità del silenziamento dell'espressione. Pertanto, l'introduzione in organismi del gene *qde-1*, o di geni omologhi, può costituire un sistema per reprimere più efficacemente funzioni geniche. In particolare, questo approccio è particolarmente utile nelle piante dove la co-soppressione viene comunemente usata per il "knock out" di funzioni geniche. Sempre nelle piante il potenziamento del silenziamento genico può essere usato allo scopo di ottenere linee resistenti a virus patogeni, per introduzione di transgeni codificanti sequenze virali, al fine di ottenere l'inibizione dell'espressione del virus stesso (Flavell et al., 1994).

Analoghe applicazioni sono trasponibili in ambito animale, dove diverse osservazioni indicano che fenomeni di silenziamento possono

determinare l'impossibilità di un'adeguata espressione dei transgeni introdotti (Garrick et al. 1998).

Al contrario vi sono casi in cui si desidera evitare e/o ridurre il silenziamento genico, per esempio quando s'intende sovraesprimere un transgene. E' noto che la co-soppressione é strettamente correlata sia con la presenza di un alto numero di copie di transgene, che con l'alta espressione dei transgeni. Questa correlazione può rendere difficile ottenere organismi transgenici che esprimono ad alti livelli un transgene, in quanto più é alta l'espressione e/o il numero di copie più é probabile scatenare risposte di silenziamento. Come già accennato, analoghi meccanismi di inattivazione genica, dipendenti da un alto numero di copie, sono stati evidenziati anche in sistemi animali. In questi casi linee di piante o linee animali, incapaci totalmente o parzialmente di silenziare, costituiscono l'ideale recipiente in cui sovraesprimere il transgene d'interesse. L'invenzione può trovare applicazioni in questo ambito usando diversi approcci:

A) Identificazione e produzione di linee mutanti nei geni omologhi al gene *qde-1* in piante, animali e funghi.

La conoscenza del gene *qde-1* di *Neurospora*, essenziale per il meccanismo di silenziamento può consentire l'isolamento di linee mutanti in altri organismi, negli omologhi del gene *qde-1*. Ad esempio, attraverso amplificazioni usando primers degenerati, disegnati sulle regioni conservate del gene *qde-1*, si possono identificare linee mutanti nei geni omologhi, analizzando banche di mutanti per inserzione già disponibili per diverse specie vegetali. Nei funghi e negli animali tali

mutanti possono essere ottenuti, una volta identificato il gene omologo, sfruttando le classiche tecniche di 'gene disruption' che fanno uso della ricombinazione omologa.

B) Riduzione dell'espressione del gene *qde-1* .

Altre strategie, per ottenere linee incapaci di silenziare, prevedono l'uso o del gene *qde-1* di *Neurospora* o di geni omologhi ad esso. Il gene *qde-1* od omologhi possono essere introdotti in vettori d'espressione opportuni ed espressi in direzione antisenso per bloccare l'espressione dei geni endogeni residenti. Alternativamente, porzioni di *qde-1* od omologhi possono essere sovraesprese, nell'intento di ottenere un effetto dominante negativo, che blocchi la funzione dei geni *qde-1* endogeni.

Gli autori della presente invenzione hanno clonato e caratterizzato il gene *qde-1* di *Neurospora crassa*. L'analisi della sequenza del gene *qde-1* ha rivelato una regione con una significativa omologia con la sequenza di una RNA polimerasi RNA-dipendente isolata da pomodoro per la quale è stato suggerito, ma non dimostrato, un coinvolgimento nel meccanismo di co-soppressione (Schiebel et al., 1998).

Gli autori dell'invenzione hanno per la prima volta dimostrato che un gene codificante per una RNA polimerasi RNA-dipendente è coinvolto nel silenziamento genico indotto da transgeni. Quindi, per la prima volta, è dimostrato che un gene appartenente alla famiglia delle RNA polimerasi RNA-dipendenti è un componente essenziale del meccanismo di inattivazione delle sequenze ripetute.

ING. DANIELLO ZANARDI ROMA SpA



Nell'ambito dell'invenzione il riferimento a percentuali di omologia va inteso come percentuale di similarità, cioè numero di residui identici + numero di residui conservati, rispetto al totale di residui della sequenza considerata.

Forma pertanto oggetto della presente invenzione una sequenza nucleotidica codificante per una proteina caratterizzata dal fatto di avere un'attività di silenziamento e di comprendere un dominio di RNA polimerasi RNA-dipendente, in cui il dominio è omologo almeno per il 30% con la sequenza amminoacidica dall'amminoacido 710 all'amminoacido 1282 di SEQ ID No.1. Preferibilmente il dominio è omologo almeno per il 40% con la sequenza amminoacidica dall'amminoacido 710 all'amminoacido 1282 di SEQ ID No.1. Più preferibilmente il dominio è omologo almeno per il 50% con la sequenza amminoacidica dall'amminoacido 710 all'amminoacido 1282 di SEQ ID No.1. Ancora più preferibilmente il dominio comprende la sequenza amminoacidica dall'amminoacido 710 all'amminoacido 1282 di SEQ ID No.1. In una forma particolare di attuazione, la sequenza nucleotidica codifica per una proteina avente la sequenza amminoacidica di SEQ ID No.1, o porzioni funzionali di essa. Ancora più preferibilmente la sequenza nucleotidica è la sequenza nucleotidica di SEQ ID No.1 o la sua sequenza complementare.

Ulteriore oggetto dell'invenzione è un vettore di espressione comprendente sotto il controllo di un promotore esprimibile in batteri la sequenza nucleotidica dell'invenzione. Gli esperti del settore realizzeranno che qualsiasi plasmide in grado di esprimere in maniera

corretta e efficace la proteina dell'invenzione in batteri può essere utilizzato e rientra nell'ambito dell'invenzione stessa.

Ulteriore oggetto dell'invenzione è un vettore di espressione comprendente sotto il controllo di un promotore esprimibile in piante o in parti di esse la sequenza nucleotidica dell'invenzione, in orientamento senso o antisenso. Gli esperti del settore realizzeranno che qualsiasi plasmide in grado di esprimere in maniera corretta e efficace la proteina dell'invenzione in piante o in parti di esse può essere utilizzato e rientra nell'ambito dell'invenzione stessa.

Ulteriore oggetto dell'invenzione è un vettore di espressione comprendente sotto il controllo di un promotore esprimibile in funghi la sequenza nucleotidica dell'invenzione, in orientamento senso o antisenso. Gli esperti del settore realizzeranno che qualsiasi plasmide in grado di esprimere in maniera corretta e efficace la proteina dell'invenzione in funghi può essere utilizzato e rientra nell'ambito dell'invenzione stessa.

Ulteriore oggetto dell'invenzione è un vettore di espressione comprendente sotto il controllo di un promotore esprimibile in animali la sequenza nucleotidica dell'invenzione, in orientamento senso o antisenso. Gli esperti del settore realizzeranno che qualsiasi plasmide in grado di esprimere in maniera corretta e efficace la proteina dell'invenzione in animali può essere utilizzato e rientra nell'ambito dell'invenzione stessa.

E' ulteriore oggetto dell'invenzione un organismo procariotico trasformato con il vettore di espressione in batteri dell'invenzione.

ING. BATTISTO & ZANARDI ROMA SPA

E' ulteriore oggetto dell'invenzione una pianta o parti di essa trasformata con il vettore di espressione in piante dell'invenzione.

E' ulteriore oggetto dell'invenzione una pianta mutata nella sequenza nucleotidica dell'invenzione con un'attività di silenziamento ridotta o inibita.

E' ulteriore oggetto dell'invenzione un fungo trasformato con il vettore di espressione in funghi dell'invenzione.

E' ulteriore oggetto dell'invenzione un fungo mutato nella sequenza nucleotidica dell'invenzione con un'attività di silenziamento ridotta o inibita.

E' ulteriore oggetto dell'invenzione un animale non umano trasformato con il vettore di espressione in animali dell'invenzione.

E' ulteriore oggetto dell'invenzione un animale non umano mutato nella sequenza nucleotidica dell'invenzione con un'attività di silenziamento ridotta o inibita.

Costituisce un ulteriore oggetto dell'invenzione una proteina caratterizzata dal fatto di avere un'attività di silenziamento e di comprendere un dominio RNA polimerasi RNA-dipendente, in cui il dominio è omologo almeno per il 30% con la sequenza amminoacidica dall'amminoacido 710 all'amminoacido 1282 di SEQ ID No.1. Preferibilmente il dominio è omologo almeno per il 40% con la sequenza amminoacidica dall'amminoacido 710 all'amminoacido 1282 di SEQ ID No.1. Più preferibilmente il dominio è omologo almeno per il 50% con la sequenza amminoacidica dall'amminoacido 710 all'amminoacido 1282 di SEQ ID No.1. Ancora più preferibilmente il

dominio è la sequenza amminoacidica dall'amminoacido 710 all'amminoacido 1282 di SEQ ID No.1. In una forma particolare di attuazione, la proteina comprende la sequenza amminoacidica di SEQ ID No.1, o porzioni funzionali di essa.

Rientra nell'ambito della presente invenzione l'uso della sequenza nucleotidica dell'invenzione stessa per la modulazione del silenziamento genico in piante, animali e funghi.

La presente invenzione verrà ora descritta in suoi esempi non limitativi con riferimento alle seguenti figure:

La figura 1 illustra il ripristino dell'espressione di *al-1* nel ceppo mutante inserzionale 107. L'RNA totale è stato estratto da miceli prelevati dopo 10 minuti di induzione alla luce, di un ceppo silenziato per *al-1* (6XW), di un ceppo non trasformato selvatico (WT) e del ceppo mutante 107. Per la ibridazione, è stata utilizzata una sonda specifica per *al-1*. La normalizzazione utilizzando una sonda specifica *al-2* è mostrata al di sotto.

La figura 2 illustra la organizzazione genomica del gene *qde-1*.

a) Sono rappresentati i due cosmidi (56G11 e 40H7) in grado di complementare i mutanti *qde-1*. Il riquadro chiaro nel cosmide 40H7 rappresenta le sequenze dal vettore cosmidico. Una mappa di restrizione del frammento di 7.9 Kb ottenuto con EcoRI, da 40H7, contenente *qde-1* è mostrato: E (EcoRI), P(PstI), B(BglII). Il riquadro scuro rappresenta la ORF identificata all'interno del frammento di 7.9 Kb di EcoRI. I plasmidi pDX e pSX contenenti i frammenti di DNA subclonati nei siti XbaI (X) e EcoRI (E) sono anche mostrati. b) Analisi

INC. BARTOLO ZANARDI ROMA SPA



per Southern dei ceppi 107 e WT. Il DNA genomico è digerito con BglI e NaeI. Nel diagramma al di sotto è mostrata la sonda di DNA utilizzata per la ibridazione e i frammenti di restrizione BglI/NaeI (B/N) attesi. Il triangolo rappresenta il sito di integrazione nel ceppo 107 che provoca una scomparsa del frammento di restrizione di 1.0 Kb.

La figura 3 illustra l'espressione di *qde-1* nel ceppo mutante inserzionale 107, nel ceppo selvatico non trasformato (WT) e nel ceppo silenziato per *al-1* (6XW). L'RNA totale è stato ibridato con una sonda specifica per *qde-1*. La quantità di RNA caricata su gel è mostrata di sotto.

La figura 4 rappresenta la sequenza amminoacidica dedotta dal gene *qde-1*. La sottolineatura indica il dominio conservato RdRP come mostrato nell'allineamento in figura 5.

La figura 5 mostra un allineamento di sequenza della proteina QDE1 con altri polipeptidi dal database SwissProtien sequence: ORF di *C. elegans* Z488334 (*eleg1*), ORF di *S. pombe* Z98553 (*pom*), ORF di *A. thaliana* AF080120 (*araB*) e la RNA polimerasi RNA dipendente di pomodoro Y10403 (RdRP). I residui identici sono evidenziati in nero, mentre le sostituzioni conservative sono mostrate in grigio.

MATERIALI E METODI

Cepi, condizioni di crescita e trasformazione

La metodologia e l'analisi degli eterocarion in *Neurospora crassa* è sostanzialmente la stessa di quella descritta in (Davis and De Sevres, 1970). Gli sferoplasti sono stati preparati secondo il protocollo di Vollmer e Yasnofsky (1997). Il ceppo 107 è stato isolato come

segue: un ceppo silenziato per *qde-1*, denominato 6xw già descritto (Cogoni and Macino, 1997) è stato trasformato con pMXY2 che contiene il gene della beta-tubulina per la resistenza al benomile, che funziona come marcatore selezionabile dominante in *N.crassa* (Staben et al., 1989). I trasformanti resistenti al benilato sono stati selezionati in base alla produzione di carotenoidi per visualizzazione del colore dei conidi: i selvatici erano arancioni, mentre i trasformanti da bianco a giallo erano indicativi di un'attività di silenziamento.

Plasmidi e genoteche

Il gene genomico *qde-1* è stato isolato da una genoteca cosmidica di *N.crassa* (Cabibbo et al., 1991). La subclonazione dei frammenti di restrizione dai cloni della genoteca è stata effettuata nel plasmide pBSK. Conseguentemente, i subcloni sono stati utilizzati in esperimenti di cotrasformazione utilizzando pMXY2 oppure pES200 (contenente la resistenza per la igromicina).

Ibridazioni Southern e Northern

Il DNA cromosomale è stato preparato come descritto da Morelli et al (1993). Dopo digestione, il DNA genomico è stato trasferito come descritto in Maniatis et al (1982). Le sonde sono state marcate per innesco casuale (Boehringer). L'RNA è stato sottoposto ad elettroforesi su gel di agarosio, trasferito su membrane Hybond-N e sondato.

Analisi e sequenziamento del DNA

La sequenza nucleotidica di *qde-1* è stata determinata per entrambe le eliche utilizzando la TAQ FS polimerasi e il metodo

fluorescente, e analizzata utilizzando un apparato automatico Applied Biosystems 373A: Le sequenze nucleotidiche e quella amminoacidica derivata sono state analizzate utilizzando il programma MacMolly Tetra. Un paragone delle proteine è stato effettuato con le ricerche BLASTP. L'algoritmo ClustalW è stato utilizzato per l'allineamento.

Risultati

Per clonare i geni *qde*, è stata utilizzata una mutagenesi inserzionale su un ceppo transgenico per *al-1* (6XW) che mostra un fenotipo albino (bianco) come conseguenza di un silenziamento post-trascrizionale del gene endogeno *al-1*: Tra i 100.000 trasformanti inserzionali indipendenti, un ceppo (107) ha mostrato una reversione del silenziamento genico visibile come recupero di un fenotipo selvatico arancione. Il fenotipo selvatico arancione del ceppo 107 è dovuto al ripristino dell'espressione del mRNA di *al-1*, come dimostrato da un'analisi Northern (si veda figura 1). Inoltre, un'analisi di eterocarion ha rivelato che la mutazione è recessiva e agisce in *trans*. L'analisi di eterocarion ha anche reso possibile assegnare il mutante nel ceppo 107 ad uno dei tre gruppi di complementazione *qde* già identificati (Cogoni and Macino 1997). Il restauro di un fenotipo silenziato per *al-1* avviene in eterocarion con mutanti *qde-2* e *qde-3*. Non è possibile complementare con mutanti *qde-1* (Tabella 1), indicando che il ceppo 107 è mutato nel gene *qde-1*.

Tabella 1 - Il ceppo 107 è mutato nel gene *qde-1*

		Ceppi mutanti <i>qde</i> utilizzati in eterocarion specifici					
	107	M17	M18	M10	M11	M7	M20
107	WT	AL	AL	AL	AL	WT	WT
M17		WT	WT	AL	AL	AL	AL
M18			WT	AL	AL	AL	AL
M10				WT	WT	AL	AL
M11					WT	AL	AL
M7						WT	WT
M20							WT

WT= eterocarion con fenotipo selvatico per l'accumulo di carotenoidi (arancione); AL= eterocarion con fenotipo albino in cui il silenziamento del gene *al-1* è ripristinato. I mutanti *qde* sono stati descritti in Cogoni and Macino (1997): M17 e M18 sono mutanti *qde-3*; M10 e M11 sono mutanti *qde-2*; M7 e M20 sono mutanti *qde-1*. I fenotipi degli eterocarion sono stati esaminati per ispezione visiva dopo sette giorni di crescita alla luce.

Per isolare il gene *qde-1*, il plasmide "tagging" è stato recuperato dal ceppo 107 con una procedura di liberazione del plasmide, utilizzando l'enzima di restrizione *Bgl*II che ha un singolo sito di restrizione all'interno del plasmide tagging. Un plasmide, pCR4 è stato recuperato dopo riligazione del DNA cromosomale a seguito di restrizione con *Bgl*II (si veda figura 2a). Il DNA cromosomale che fiancheggia il sito di inserzione è stato isolato utilizzando gli enzimi *Bgl*II e *Pst*I e il frammento di restrizione risultante è stato utilizzato

ING. BARZANO & ZANARDO ROMA S.p.A.



come sonda su una genoteca cosmitica di *N.crassa*. Due cosmidi ibridizzanti positivi 56G11 e 40H7 sono stati introdotti con un esperimento di trasformazione nel ceppo 107 e nel ceppo mutante per *qde-1* M7 indotto per UV. Entrambi i cosmidi sono in grado di complementare i due mutanti *qde-1* saggiati. Il ripristino del silenziamento del gene *al-1* che conduce alla comparsa di un fenotipo bianco, indica che entrambi i cosmidi contengono un gene *qde-1* funzionale. Utilizzando lo stesso DNA fiancheggiante come sonda, due frammenti di *EcoRI* da 40H7 e da 56G11, rispettivamente di 7.9 e 10.0 Kb, sono stati identificati e subclonati (si veda figura 2a). Entrambi i frammenti di *EcoRI* (plasmidi P79E e p10E) sono in grado di complementare il ceppo *qde-1* (107 e M7), ma non i ceppi mutanti in *qde-2* (M10) e *qde-3* (M17) (Tabella 2) indicando che il gene funzionale *qde-1* è contenuto nel frammento *EcoRI* di 7.9 Kp.

Tabella 2 Complementazione con il gene *qde-1*

Plasmidi	ceppi mutanti qde utilizzati			
	107	M7(<i>qde-1</i>)	M10(<i>qde-2</i>)	M17(<i>qde-3</i>)
P79E	58/200(29%)	48/200(24%)	0/200(0%)	0/200(0%)
P10E	51/100(25%)	25/100(25%)	0/200(0%)	0/200(0%)
PSX	0/200(0%)	0/200(0%)	-	-
PDX	0/200(0%)	0/200(0%)	0/200(0%)	0/200(0%)

La frequenza di complementazione è riportata come percentuale dei ceppi trasformanti che mostrano un fenotipo albino rispetto al numero totale di trasformanti. Il plasmide pMXY2 è stato usato come controllo negativo.

Il fatto che il gene *qde-1* sia in grado di ripristinare il silenziamento del gene *al-1* solo nei mutanti corrispondenti, esclude la possibilità che il frammento di DNA clonato sia in grado di riattivare il silenziamento del gene *al-1*, a prescindere dal gruppo di complementazione *qde*. Inoltre, il ceppo M7 trasformato con il frammento di 7.9 Kb *EcoRI* conferisce una capacità di silenziare un altro gene carotenogenico *al-2*, quando introdotto per trasformazione (dati non mostrati). Il frammento *EcoRI* di 7.9 Kb è stato ulteriormente subclonato utilizzando il sito *XbaI* (si veda figura 2a) che taglia al centro del frammento *EcoRI*. Entrambi i frammenti *XbaI/EcoRI* (plasmidi pSX e pDX) non erano in grado di complementare i ceppi 107 e M7 (si veda Tabella 2), suggerendo che il sito *XbaI* è probabilmente localizzato all'interno del gene *qde-1*.

L'intera regione del frammento *EcoRI* che comprende il gene *qde-1* putativo e le regioni adiacenti sono state sequenziate, rivelando una cornice di lettura aperta (ORF) di 4206 bp che codifica per una proteina putativa di 1402 amminoacidi. Due risultati diversi indicano che questa ORF è coincidente con il gene *qde-1*. In primo luogo, il sito di restrizione *XbaI* utilizzato per subclonare il frammento di 7.9 Kb *EcoRI* è localizzato al centro della ORF (figura 2a), ed è pertanto consistente con il risultato che entrambi i frammenti *KbaI/EcoRI* non siano in grado di complementare la mutazione *qde-1* (Tabella 2). In secondo luogo, per analisi Southern (si veda figura 2b) il sito di inserzione del plasmide tagging nel ceppo 107 è stato mappato all'interno dei siti di restrizione *BglII* e *NaeI* all'interno della ORF. Non è

stata rivelata alcuna alterazione di dimensione delle regioni fiancheggianti, escludendo la possibilità che le delezioni comprendano altre ORF all'interno della regione di 7.9 Kb di EcoRI. I cDNA, sintetizzati per PCR inversa (RT-PCR), hanno rivelato una colinearità con il DNA genomico indicando che non è presente alcun introne. L'espressione del gene *qde-1* è stata analizzata per analisi Northern utilizzando una sonda che comprende la ORF di *qde-1* (figura 2a). E' stato così rivelato un trascritto di circa 5000 nt, e il livello basale di mRNA di *qde-1* è risultato essere due volte più alto in un ceppo silenziato per *al-1* (6XW) rispetto ad un ceppo WT non trasformato (si veda figura 3). Inoltre, l'mRNA di intera lunghezza di *qde-1* non è rivelabile nel ceppo mutante inserzionale 107 di *qde-1*, dove invece una banda più piccola è presente suggerendo che i trascritti troncati di *qde-1* sono prodotti come conseguenza dell'integrazione. Il fatto che l'espressione genica di *qde-1* sia aumentata in maniera specifica in un ceppo silenziato indica l'esistenza di un meccanismo regolativo in grado di attivare componenti cellulari del macchinario di silenziamento nei ceppi trasgenici.

La proteina QDE-1 dedotta dalla sequenza nucleotidica è composta di 1402 amminoacidi (si veda figura 4) ed ha un peso molecolare di 158.004 Da ed un pI statistico di 8.0. La proteina QDE-1 non contiene un peptide segnale o un dominio transmembrana, indicando che è probabilmente una proteina intracellulare. Inoltre, il plot di idropatia indica che QDE-1 è una proteina solubile. Una ricerca BLAST ha mostrato che *qde-1* ha una omologia statisticamente

significativa con proteine ipotetiche da diversi altri organismi comprendenti: due ORF di *C.elegans* (numeri di accesso EMBL Z4834 e Z78419) con valori attesi (valore E) di $2e-16$ e $9e-10$ rispettivamente; una ORF di *S.pombe* (numero di accesso EMBL Z98553) con un valore E $3e-13$; quattro ORF di *A.thaliana* (numeri di accesso EMBL AF080120, AC005169 comprendente tre ORF alla stessa localizzazione cromosomale) e con un valore E $8e-15$, $7e-06$, $4e-05$, $5e-02$, rispettivamente. Infine è stata trovata una omologia significativa (valore E $2e-17$) con una proteina putativa codificata da un cDNA di pomodoro (numero di accesso EMBL Y10403). L'omologia trovata non si estende per l'intera proteina, ma è ristretta ad una porzione di 570 amminoacidi, dall'amminoacido 710 all'amminoacido 1282, che definisce un dominio conservato (si veda figura 5). Tra le proteine putative omologhe identificate, solo quella derivata dalla sequenza del cDNA di pomodoro è stata caratterizzata in maniera funzionale come una RNA polimerasi RNA dipendente (RdRP, 9).

ING. BARTOLO & ZANARDI ROMA S.p.A.



SEQUENCE LISTING

<110> Universita' degli Studi di Roma "La Sapienza"

<120> Isolamento e caratterizzazione di un gene per il
silenziamento genico da N. crassa e suoi usi

<160> 2

<170> PatentIn Ver. 2.0

<210> 1

<211> 8045

<212> DNA

<213> Neurospora crassa

<221> CDS

<222> (2447)..(6655)

<400> 1

```

ggtaccgggc ccccccctcga ggtcgacggt atcgataagc ttgatatcga attctgaaaa 60
gtcctaaagt ctctgtctc tcattgctagg tagccctact gcgtgtgtag ttagtagggt 120
ccgaatagggt tcgtacgctg tcccccttag ggggctggtt taacagggag acaggggtaca 180
gtctctcggg aacagagacg accaagcttt gttgtgtgac atcatcttga gaccattcgt 240
ctgaggggtcc atacaacccat acattgctgtg tagagagtgg tgggaatatt agaggttact 300
ttgagagttt ctgaagagat tgagactgta ggtacgttca ttatcgggtg cgacgagaat 360
gggcacgcat gattcgacat ttgcaatttg ttggttggca cgcacatcga tagtgtacac 420
atgtgtcggc ctgagactgt ccaatggacc aatggtaagg cgtgagcggc gattggggtc 480
aacttgagggt taattgtacac agtacttgag gctgtagcgg gcagcgggaag catccgtcct 540
cgaatggaca ttcccaagcc agccaacact ggtgtccatt taaaagcgtt ttctccatgg 600
gggaacaatg gccggcgtag tatctgtaag aggacctaga ggacgggtatt agaacatttg 660
ggtgatgtat ctagatcacc gatgtgtccc tgaatagtcc atgaatagtc tgggatgagt 720
ccgagatcag tccgggatgt gtcaccgata tatccatgac ttatctctga ttatcccag 780
cgagacccca atcagtatga ccgacagcga caccatcacc ggcgccagcg tatgagatcc 840
tgaacatatg ttaggcatat gtatggcata tgtatgcata gagcatgcac cacatatctg 900
agagctgcct gtgactacct agaggatatg agaggtaatg tcacgtacct tacctctacc 960
aaatgtatcc cgcagtcact cgtattcccg ctgtagtgtg cactagaact agccccgcct 1020
tttggatcct ggagtgtccc cattccacgc acttgccaag aaccacaacc tcgcaaagcg 1080
cgctgacagg gcaccgtgct gcctggacct ttgcgggcgg ttcttcttg accggcgcca 1140
ggcagcagac cggttggttc ttgttctca aggtgacgaa ggcgattaaa aaatccaaga 1200
caaccagatc ctggaacctt tgatcatccc tcgtttactc agacgaaaaa actaaaaaac 1260
ttcaaagtcg ttgtcagatt aattgaaccg ttaaaacagg acattgtccgc cctttcgtct 1320
ctcatgccct cgaaacctca cgtcttccat ctctcgtccg atccccccct cccacgacc 1380
ctgattcagc agcgaactga cctgtcactt tactgtggac tctcctaga gtagggctgg 1440

```

ING. BAZZANO ZAMANO ROMA SPA

ccatccccgg gccccctgca gcttctgcaa gtcaaagcga tggaaggatg gccgtccagt 1500
 ggggtgcggca tacagccatt tccaaaatgg tctcgtcgca tgcggaagaa aaaaaatgga 1560
 aggctacgag cctgcgacac ctgtacggtc actgcggtc cccttcacca ccaccagtga 1620
 tatcatcacc cataccatca gggccactac cgccgcgctc tctcgaacat cgggctgaat 1680
 gtgaccaatt gcgccgacag cgggcttttc gatcctgtgg gcagcttctt ccaccggttt 1740
 gcgctgatct gggagcgtat cgagctggat tgggactggt taccactcga atcacttgac 1800
 ggctcggcgc gattctatgg ccattccaccg cgattgatcg gtcccttcac tgtttcctgc 1860
 aacctactct cccgaccaga tcagacgagg ctttatgcc agttccgggc gttgccaatt 1920
 atagccacag cacgaagctc gccaagtgc ttcgaaatgc tcaaaactgt tcacacactt 1980
 tgcacgaaca ggaccaccac gaaaaaccac cttggccgc tcgcgcgaac ccgtgtcttt 2040
 gaactcgatc agccccaatc ccaaaccagg tcgcctttcc ccgctgcgcc tgcttgctcc 2100
 gacttcgtga agcaaacaag actggcggct atggccaagg ccattccat gacaattggt 2160
 taatcgaggt caaaaactac aagagatcaa cgccgcaaag atatacttgt cgttcacccc 2220
 gacaaccgca aagccttgaa gcgcaaagcc acagtgggtg tgtttacttc cgcacactta 2280
 tcgaccctta cgcccaagac gctcccgctt ggtagaagct cgagtcgggt aatcaggat 2340
 atctcacagt cgcttgctc tccttttcgg tggtagctgc ttttatcatg ccttatcgct 2400
 aaccgggttg tatactctac tccaggcccg atctcccacg gcacaa atg aac cct 2455

Met Asn Pro

1

att act cct agg aag agg aat agc ccc gtc gag gaa atc ata aac cgg 2503
 Ile Thr Pro Arg Lys Arg Asn Ser Pro Val Glu Glu Ile Ile Asn Arg
 5 10 15
 ctc aat aac gac tac aac ctg ggc ctc cag tgt gtc gca gac aca act 2551
 Leu Asn Asn Asp Tyr Asn Leu Gly Leu Gln Cys Val Ala Asp Thr Thr
 20 25 30 35
 ctc acc ccc cac cgc cgg aag gag ctg gcc gag agt gac gag gat ttc 2599
 Leu Thr Pro His Arg Arg Lys Glu Leu Ala Glu Ser Asp Glu Asp Phe
 40 45 50
 ggt cgc cat gac aag atc tac aga gcc ctg aac ttt ctc tac tgg cgg 2647
 Gly Arg His Asp Lys Ile Tyr Arg Ala Leu Asn Phe Leu Tyr Trp Arg
 55 60 65
 aag gat gac tcc ctg aac cag gca gaa gcc aac ttc ttc atc gag gcc 2695
 Lys Asp Asp Ser Leu Asn Gln Ala Glu Ala Asn Phe Phe Ile Glu Ala
 70 75 80
 aaa gct gcg agc tcg aac tgg gtg ccc aaa gcc cac gcc gac cct gac 2743
 Lys Ala Ala Ser Ser Asn Trp Val Pro Lys Ala His Ala Asp Pro Asp
 85 90 95
 acg ctt ccg tgg tcc aag gaa cct ccc cgc gcc gct act gcc ggc caa 2791

ING. BATAK & ZANARDI 1974 834

Thr Leu Pro Trp Ser Lys Glu Pro Pro Arg Ala Ala Thr Ala Gly Gln	
100	105 110 115
caa tgg gca ttg cag act gtg ttg ctc gag gtg ctt aat agg ttt atg	2839
Gln Trp Ala Leu Gln Thr Val Leu Leu Glu Val Leu Asn Arg Phe Met	
120	125 130
cca cct ccc aat aac aca cca ggt cga acg ttt ggc aga act cta agc	2887
Pro Pro Pro Asn Asn Thr Pro Gly Arg Thr Phe Gly Arg Thr Leu Ser	
135	140 145
ggc cca agt ggc ctg agc cgc cca acc tct acc aac acc aaa cgc aag	2935
Gly Pro Ser Gly Leu Ser Arg Pro Thr Ser Thr Asn Thr Lys Arg Lys	
150	155 160
gat gag ccc gcc aat gtc act ttc gct gat ccg ccc aaa cgc tcg ttg	2983
Asp Glu Pro Ala Asn Val Thr Phe Ala Asp Pro Pro Lys Arg Ser Leu	
165	170 175
act cgc tct gcc aca ggt cct cct att cac ggc gcg gcg ata ccc cta	3031
Thr Arg Ser Ala Thr Gly Pro Pro Ile His Gly Ala Ala Ile Pro Leu	
180	185 190 195
aag ttc ccc gat cca gtg aat acc ggt tcc aaa cga cca tct ctc gag	3079
Lys Phe Pro Asp Pro Val Asn Thr Gly Ser Lys Arg Pro Ser Leu Glu	
200	205 210
agt gag aat ctc aat cag tgc acc aag cgg gcc aag ggc aag ctg tct	3127
Ser Glu Asn Leu Asn Gln Cys Thr Lys Arg Ala Lys Gly Lys Leu Ser	
215	220 225
gat aat gtt gcc gct gcc gcc gcc ccg ccc gtg cct att gcg agc gct	3175
Asp Asn Val Ala Ala Ala Ala Ala Pro Pro Val Pro Ile Ala Ser Ala	
230	235 240
ttg gac aag gta ccg act cga agg cat gcc aat acg aga gat ccc acg	3223
Leu Asp Lys Val Pro Thr Arg Arg His Ala Asn Thr Arg Asp Pro Thr	
245	250 255
gcg aca ggt cat aga cga gcg gac cag gtg gat tcc ttt gat aca tct	3271
Ala Thr Gly His Arg Arg Ala Asp Gln Val Asp Ser Phe Asp Thr Ser	
260	265 270 275
caa ggc act tcc tat ggt tcg agt gtc ttc agc gct tgc cgt cac aat	3319
Gln Gly Thr Ser Tyr Gly Ser Ser Val Phe Ser Ala Cys Arg His Asn	
280	285 290
cag agc act acc cag agt agt ttt gag gct cct cct tca cag ccc aga	3367
Gln Ser Thr Thr Gln Ser Ser Phe Glu Ala Pro Pro Ser Gln Pro Arg	
295	300 305

gag aag cgg cct gtg gat gcc acg gtc ttt gag gct gga cac ttg att 3415
 Glu Lys Arg Pro Val Asp Ala Thr Val Phe Glu Ala Gly His Leu Ile
 310 315 320
 gag tct cct agc aaa gga aga aca acc aag tcc cac ata gat aac cag 3463
 Glu Ser Pro Ser Lys Gly Arg Thr Thr Lys Ser His Ile Asp Asn Gln
 325 330 335
 ccc ctt tca tcg tct tcc cag ggt gaa act tcg ttc agc act tac tat 3511
 Pro Leu Ser Ser Ser Ser Gln Gly Glu Thr Ser Phe Ser Thr Tyr Tyr
 340 345 350 355
 gag tcg ttt cca agt tcc ggc ggc gag ggc gca att ccc gag ccg agt 3559
 Glu Ser Phe Pro Ser Ser Gly Gly Glu Gly Ala Ile Pro Glu Pro Ser
 360 365 370
 cgc tca aat gga ctg gct cgg agc gaa gaa agc gct cga tct cag gtt 3607
 Arg Ser Asn Gly Leu Ala Arg Ser Glu Glu Ser Ala Arg Ser Gln Val
 375 380 385
 caa gtt cat gcc ccg gtg gtt gca gct cgg ctg aga aat att tgg ccg 3655
 Gln Val His Ala Pro Val Val Ala Ala Arg Leu Arg Asn Ile Trp Pro
 390 395 400
 aaa ttt ccc aaa tgg cta cac gaa gct cct ctc gct gtt gca tgg gaa 3703
 Lys Phe Pro Lys Trp Leu His Glu Ala Pro Leu Ala Val Ala Trp Glu
 405 410 415
 gtt acc aga ctc ttt atg cac tgc aaa gta gac ttg gaa gac gag agc 3751
 Val Thr Arg Leu Phe Met His Cys Lys Val Asp Leu Glu Asp Glu Ser
 420 425 430 435
 ctg ggc cta aag tac gac cct tcc tgg tct acc gcg cgc gat gtc aca 3799
 Leu Gly Leu Lys Tyr Asp Pro Ser Trp Ser Thr Ala Arg Asp Val Thr
 440 445 450
 gat atc tgg aag act ctc tac cgg ctt gat gct ttc cgt ggt aaa ccc 3847
 Asp Ile Trp Lys Thr Leu Tyr Arg Leu Asp Ala Phe Arg Gly Lys Pro
 455 460 465
 ttt cca gaa aag ccg ccc aac gac gtg ttc gtg acc gca atg acg ggc 3895
 Phe Pro Glu Lys Pro Pro Asn Asp Val Phe Val Thr Ala Met Thr Gly
 470 475 480
 aac ttt gag agc aaa ggt agt gcc gtt gtt ctc tct gct gtt cta gac 3943
 Asn Phe Glu Ser Lys Gly Ser Ala Val Val Leu Ser Ala Val Leu Asp
 485 490 495
 tac aat ccg gac aac tcg cct act gcg ccc ctt tac ctt gtg aag ctg 3991
 Tyr Asn Pro Asp Asn Ser Pro Thr Ala Pro Leu Tyr Leu Val Lys Leu

ING. BARZANO & ZANARDO PARRA



500	505	510	515	
aag ccg ctc atg ttc gag cag ggc tgt cga ctc acc cgt cgg ttc ggt				4039
Lys Pro Leu Met Phe Glu Gln Gly Cys Arg Leu Thr Arg Arg Phe Gly				
	520	525	530	
cct gat agg ttt ttc gag atc ctt ata ccg tcg cct acg agc acc agc				4087
Pro Asp Arg Phe Phe Glu Ile Leu Ile Pro Ser Pro Thr Ser Thr Ser				
	535	540	545	
cca agt gta ccg ccg gtg gtc agc aaa caa cca ggt gcg gtc gaa gaa				4135
Pro Ser Val Pro Pro Val Val Ser Lys Gln Pro Gly Ala Val Glu Glu				
	550	555	560	
gtc atc cag tgg ctc acg atg ggg caa cat tct ctt gta ggc cgc caa				4183
Val Ile Gln Trp Leu Thr Met Gly Gln His Ser Leu Val Gly Arg Gln				
	565	570	575	
tgg cgc gct ttc ttc gcc aaa gat gcc gga tac agg aaa cct ctc agg				4231
Trp Arg Ala Phe Phe Ala Lys Asp Ala Gly Tyr Arg Lys Pro Leu Arg				
580	585	590	595	
gag ttc cag ctc cgc gcc gag gac ccg aaa ccc atc atc aag gag aga				4279
Glu Phe Gln Leu Arg Ala Glu Asp Pro Lys Pro Ile Ile Lys Glu Arg				
	600	605	610	
gtc cac ttc ttt gcc gag acc ggc att acg ttc cga cct gat gtg ttc				4327
Val His Phe Phe Ala Glu Thr Gly Ile Thr Phe Arg Pro Asp Val Phe				
	615	620	625	
aag acg aga tct gtc gtt ccg gca gag gaa cct gta gag caa cgg acc				4375
Lys Thr Arg Ser Val Val Pro Ala Glu Glu Pro Val Glu Gln Arg Thr				
	630	635	640	
gag ttc aaa gtt agt caa atg ctg gac tgg ctc ctg caa ctc gac aac				4423
Glu Phe Lys Val Ser Gln Met Leu Asp Trp Leu Leu Gln Leu Asp Asn				
	645	650	655	
aac act tgg cag ccg cac ctc aag ttg ttc tcc cgt atc cag ctc ggt				4471
Asn Thr Trp Gln Pro His Leu Lys Leu Phe Ser Arg Ile Gln Leu Gly				
660	665	670	675	
ctg agt aag aca tat gcc att atg aca ttg gag cct cac cag atc aga				4519
Leu Ser Lys Thr Tyr Ala Ile Met Thr Leu Glu Pro His Gln Ile Arg				
	680	685	690	
cac cac aag acc gat ctt ctt tca cct tca ggc act ggc gaa gtg atg				4567
His His Lys Thr Asp Leu Leu Ser Pro Ser Gly Thr Gly Glu Val Met				
	695	700	705	
aat gac ggt gta ggc cgc atg tcg cga agc gtg gcc aag agg ata cgc				4615

ING. BARRIO ZAMUDIO S.A.

Asn Asp Gly Val Gly Arg Met Ser Arg Ser Val Ala Lys Arg Ile Arg	
710 715 720	
gat gtt ctc ggt ttg ggt gat gtg ccc tct gct gtg caa ggg cgg ttt	4663
Asp Val Leu Gly Leu Gly Asp Val Pro Ser Ala Val Gln Gly Arg Phe	
725 730 735	
ggt tcg gcc aag gga atg tgg gtt atc gac gtt gac gac aca ggc gat	4711
Gly Ser Ala Lys Gly Met Trp Val Ile Asp Val Asp Asp Thr Gly Asp	
740 745 750 755	
gag gat tgg atc gag aca tac ccg tcc cag cgc aag tgg gaa tgc gac	4759
Glu Asp Trp Ile Glu Thr Tyr Pro Ser Gln Arg Lys Trp Glu Cys Asp	
760 765 770	
ttc gtt gat aaa cat caa cgt acc ctc gaa gtc cgg agc gtg gct tct	4807
Phe Val Asp Lys His Gln Arg Thr Leu Glu Val Arg Ser Val Ala Ser	
775 780 785	
gaa ctg aag tca gct ggt ctc aac cta cag ctg tta cct gtc ctg gaa	4855
Glu Leu Lys Ser Ala Gly Leu Asn Leu Gln Leu Leu Pro Val Leu Glu	
790 795 800	
gat aga gcc agg gac aag gtg aag atg cgc cag gca atc ggt gac cgt	4903
Asp Arg Ala Arg Asp Lys Val Lys Met Arg Gln Ala Ile Gly Asp Arg	
805 810 815	
ctt atc aac gat ttg caa cga cag ttc agc gag caa aag cat gct ttg	4951
Leu Ile Asn Asp Leu Gln Arg Gln Phe Ser Glu Gln Lys His Ala Leu	
820 825 830 835	
aat cgc cca gtg gaa ttt cgc caa tgg gtt tac gag agt tat tcc agt	4999
Asn Arg Pro Val Glu Phe Arg Gln Trp Val Tyr Glu Ser Tyr Ser Ser	
840 845 850	
cgc gca act cga gtc agc cac ggc cgt gtg cct ttt ctt gct ggg cta	5047
Arg Ala Thr Arg Val Ser His Gly Arg Val Pro Phe Leu Ala Gly Leu	
855 860 865	
cct gac agt caa gag gag aca ctg aac ttc ttg atg aac agt ggg ttc	5095
Pro Asp Ser Gln Glu Glu Thr Leu Asn Phe Leu Met Asn Ser Gly Phe	
870 875 880	
gat ccc aag aag caa aag tac ttg caa gac atc gcc tgg gat ctt caa	5143
Asp Pro Lys Lys Gln Lys Tyr Leu Gln Asp Ile Ala Trp Asp Leu Gln	
885 890 895	
aag cgg aaa tgt gac acg ttg aag tcc aag ctg aac atc cgt gtc ggt	5191
Lys Arg Lys Cys Asp Thr Leu Lys Ser Lys Leu Asn Ile Arg Val Gly	
900 905 910 915	

ING. BARZANO & ZAMBINO ROMA SPA

cga tca gca tac att tac atg att gcc gat ttc tgg ggt gtg ctt gag	5239
Arg Ser Ala Tyr Ile Tyr Met Ile Ala Asp Phe Trp Gly Val Leu Glu	
920 925 930	
gaa aat gag gtt cat gtc gga ttc tcc tca aag ttc aga gac gag gag	5287
Glu Asn Glu Val His Val Gly Phe Ser Ser Lys Phe Arg Asp Glu Glu	
935 940 945	
gag tct ttt aca ctc cta tcg gac tgt gat gtc ctc gtg gcg cga tcc	5335
Glu Ser Phe Thr Leu Leu Ser Asp Cys Asp Val Leu Val Ala Arg Ser	
950 955 960	
cca gcc cat ttc cct agt gat atc caa cgg gtt cga gca gtc ttc aag	5383
Pro Ala His Phe Pro Ser Asp Ile Gln Arg Val Arg Ala Val Phe Lys	
965 970 975	
cca gag ctc cac agt ctc aag gat gta atc atc ttc tct act aaa gga	5431
Pro Glu Leu His Ser Leu Lys Asp Val Ile Ile Phe Ser Thr Lys Gly	
980 985 990 995	
gat gta ccg ctt gct aag aag cta tct ggt gga gac tac gac ggc gat	5479
Asp Val Pro Leu Ala Lys Lys Leu Ser Gly Gly Asp Tyr Asp Gly Asp	
1000 1005 1010	
atg gcc tgg gtc tgc tgg gat ccg gag atc gtc gat ggt ttc gtc aat	5527
Met Ala Trp Val Cys Trp Asp Pro Glu Ile Val Asp Gly Phe Val Asn	
1015 1020 1025	
gcg gaa atg cct ctg gag ccc gac ctg tct agg tac cta aag aag gac	5575
Ala Glu Met Pro Leu Glu Pro Asp Leu Ser Arg Tyr Leu Lys Lys Asp	
1030 1035 1040	
aaa acg act ttc aaa caa ctt atg gcc tca cac ggc acg ggc tca gcg	5623
Lys Thr Thr Phe Lys Gln Leu Met Ala Ser His Gly Thr Gly Ser Ala	
1045 1050 1055	
gcc aaa gag cag act aca tac gat atg atc cag aag agc ttc cat ttc	5671
Ala Lys Glu Gln Thr Thr Tyr Asp Met Ile Gln Lys Ser Phe His Phe	
1060 1065 1070 1075	
gcc ctg cag ccc aac ttc ttg ggc atg tgc act aac tac aaa gaa agg	5719
Ala Leu Gln Pro Asn Phe Leu Gly Met Cys Thr Asn Tyr Lys Glu Arg	
1080 1085 1090	
ctc tgt tac atc aac aat agt gtg tct aac aag ccg gcc atc att ctt	5767
Leu Cys Tyr Ile Asn Asn Ser Val Ser Asn Lys Pro Ala Ile Ile Leu	
1095 1100 1105	
agt tca ctg gtg gga aac ctc gtc gat cag agc aag caa ggt att gtc	5815
Ser Ser Leu Val Gly Asn Leu Val Asp Gln Ser Lys Gln Gly Ile Val	

1110	1115	1120	
ttt aac gaa gca agc tgg gct caa ttg cgt agg gaa ctg ctt ggc ggt			5863
Phe Asn Glu Ala Ser Trp Ala Gln Leu Arg Arg Glu Leu Leu Gly Gly			
1125	1130	1135	
gca ttg tcc ctt cct gac cca atg tac aag agc gac agt tgg ctc ggg			5911
Ala Leu Ser Leu Pro Asp Pro Met Tyr Lys Ser Asp Ser Trp Leu Gly			
1140	1145	1150	1155
cgc gga gag cct acc cac att att gac tac ctg aaa ttc tcc atc gcc			5959
Arg Gly Glu Pro Thr His Ile Ile Asp Tyr Leu Lys Phe Ser Ile Ala			
1160	1165	1170	
agg cct gcg att gac aag gaa ctg gaa gcc ttc cac aat gcc atg aaa			6007
Arg Pro Ala Ile Asp Lys Glu Leu Glu Ala Phe His Asn Ala Met Lys			
1175	1180	1185	
gcg gcc aag gat aca gaa gac ggc gct cac ttt tgg gat ccg gat ctc			6055
Ala Ala Lys Asp Thr Glu Asp Gly Ala His Phe Trp Asp Pro Asp Leu			
1190	1195	1200	
gct tcc tac tac acg ttc ttc aag gag att agc gac aag tcg cga tcg			6103
Ala Ser Tyr Tyr Thr Phe Phe Lys Glu Ile Ser Asp Lys Ser Arg Ser			
1205	1210	1215	
tcc gca ctg cta ttc acg act ctg aag aac cgt atc ggc gaa gtc gag			6151
Ser Ala Leu Leu Phe Thr Thr Leu Lys Asn Arg Ile Gly Glu Val Glu			
1220	1225	1230	1235
aaa gaa tat ggc agg ttg gtc aaa aac aag gag atg aga gac agc aag			6199
Lys Glu Tyr Gly Arg Leu Val Lys Asn Lys Glu Met Arg Asp Ser Lys			
1240	1245	1250	
gac ccc tac cct gtc cgc gtc aac cag gtt tat gaa aaa tgg tgc gcc			6247
Asp Pro Tyr Pro Val Arg Val Asn Gln Val Tyr Glu Lys Trp Cys Ala			
1255	1260	1265	
atc acg cct gag gcg atg gac aaa tcc gga gca aat tat gat tct aag			6295
Ile Thr Pro Glu Ala Met Asp Lys Ser Gly Ala Asn Tyr Asp Ser Lys			
1270	1275	1280	
gtg atc agg ttg ctg gag ctg tcc ttc ctc gcg gac cgt gag atg aat			6343
Val Ile Arg Leu Leu Glu Leu Ser Phe Leu Ala Asp Arg Glu Met Asn			
1285	1290	1295	
aca tgg gca ttg ctg agg gct agc acg gct ttc aag ctg tac tac cac			6391
Thr Trp Ala Leu Leu Arg Ala Ser Thr Ala Phe Lys Leu Tyr Tyr His			
1300	1305	1310	1315
aag agc ccc aag ttc gtg tgg cag atg gcg ggc aga cag ctc gcg tac			6439

ING. BARTOLO & ZANARDI



Lys	Ser	Pro	Lys	Phe	Val	Trp	Gln	Met	Ala	Gly	Arg	Gln	Leu	Ala	Tyr		
1320					1325					1330							
att	aag	gcg	cag	atg	acg	agc	aga	ccc	ggt	gaa	ggc	gcc	ccg	gcg	ttg	6487	
Ile	Lys	Ala	Gln	Met	Thr	Ser	Arg	Pro	Gly	Glu	Gly	Ala	Pro	Ala	Leu		
1335					1340					1345							
atg	acc	gcg	ttc	atg	tat	gcg	ggc	ttg	atg	ccg	gat	aag	aag	ttt	acg	6535	
Met	Thr	Ala	Phe	Met	Tyr	Ala	Gly	Leu	Met	Pro	Asp	Lys	Lys	Phe	Thr		
1350					1355					1360							
aag	cag	tat	gtg	gcc	agg	ctg	gag	ggc	gat	gga	tcg	gag	tac	cct	gat	6583	
Lys	Gln	Tyr	Val	Ala	Arg	Leu	Glu	Gly	Asp	Gly	Ser	Glu	Tyr	Pro	Asp		
1365					1370					1375							
ccg	gag	gtc	tat	gaa	gtg	ctg	ggc	gat	gat	gat	ttt	gat	gga	att	ggt	6631	
Pro	Glu	Val	Tyr	Glu	Val	Leu	Gly	Asp	Asp	Asp	Phe	Asp	Gly	Ile	Gly		
1380	1385					1390					1395						
ttc	aca	ggg	aat	ggc	gat	tat	tag	tcgggttgct	tgaggggggc	cgctgcatcc						6685	
Phe	Thr	Gly	Asn	Gly	Asp	Tyr											
1400																	
ttat	tttt	tct	gatt	ttccga	gtct	tttggtg	ctgg	gaat	tg	gaat	cttctg	accat	gttcc	6745			
ctat	gggtga	gtgc	tattt	tatg	gacttc	agtt	tttgag	tttt	ctttcg	gcttt	gagat	6805					
ctgc	ctgagc	aagc	gaggcg	tggg	acattc	ggtt	ggctgc	taga	acataa	aacg	aaaatg	6865					
ggag	atgata	tgatt	cttcg	tggc	gtgata	atatt	gagat	attt	gggtgc	tgtg	ctttgt	6925					
atcc	atgtcg	cagat	attcg	ttct	acttg	ttgc	ccgcga	agcag	atttg	gtcca	acttt	6985					
caaaca	atag	tcag	caacct	tgagg	aaact	atcg	atctcg	atac	agacgc	ttaga	ataaaa	7045					
tagc	gtcctc	gtc	cccttag	cctc	cgacct	ccat	cctagg	tggag	caccc	ggg	atctcct	7105					
agact	ccttt	cagac	ctgtc	taagt	gagat	atata	gtgca	aagc	agttgg	tgtt	acactt	7165					
tggt	attaca	tcaaca	agac	aattt	gagat	agact	cattc	cctc	cggttt	tctt	ccgatt	7225					
gcgact	ctat	ttgat	gattt	ttttt	ccatc	tacc	ccctgac	aagc	ataact	ctag	caaatc	7285					
ttttt	gctga	atga	atgtag	gctc	tttggg	atttc	gaggt	ttt	gacctg	aaac	ctttga	7345					
tctttt	gttc	ttttg	ccttt	tcgt	ttagtt	ggaa	agtata	tcaact	ctac	ctgat	cacat	7405					
cgga	agcaac	aaca	acaaca	aca	acaaca	caaca	acaac	aaca	acaaca	aca	acaaca	7465					
caaca	acaac	aaca	acaaca	aca	acaaca	caaca	acaac	aaca	acaaca	aca	acaaca	7525					
caaca	acaac	aaca	acaaca	aca	acaaca	caaca	acaac	aaca	acaaca	aca	acaaca	7585					
caaca	acaac	aaca	acaaca	aca	acaaca	caaca	acaac	aaca	acaaca	aca	acaaca	7645					
cagca	acaac	aaca	acaaca	aca	acaaca	caaca	acaac	aaca	agacaa	caaca	agaca	7705					
actat	caggt	ttaac	ctcag	tacc	atttgt	aacag	accga	gcg	acctcat	agct	tatggg	7765					
aaaata	aatac	acgt	ccagag	atac	ataaca	aatac	agcaa	caaca	atgaa	gtt	caaagag	7825					
tttct	cgccc	gccg	ttccac	cacc	accagc	acc	accacca	ccc	acagcat	aaac	ctcgag	7885					
tgtat	gagtg	acg	agtacaa	cat	gaaccgt	cgac	gatccc	tata	gtgagt	cgt	attatgc	7945					

ggccgcgaat tcgatatcaa gcttatcgat accgtcgacc tcgagggggg gcccggtacc 8005
 caattcgccc tatagtgaat cgtattacaa tcccagtcac 8045

<210> 2

<211> 1402

<212> PRT

<213> Neurospora crassa

<400> 2

Met	Asn	Pro	Ile	Thr	Pro	Arg	Lys	Arg	Asn	Ser	Pro	Val	Glu	Glu	Ile
1				5					10					15	
Ile	Asn	Arg	Leu	Asn	Asn	Asp	Tyr	Asn	Leu	Gly	Leu	Gln	Cys	Val	Ala
			20					25					30		
Asp	Thr	Thr	Leu	Thr	Pro	His	Arg	Arg	Lys	Glu	Leu	Ala	Glu	Ser	Asp
		35					40					45			
Glu	Asp	Phe	Gly	Arg	His	Asp	Lys	Ile	Tyr	Arg	Ala	Leu	Asn	Phe	Leu
	50					55					60				
Tyr	Trp	Arg	Lys	Asp	Asp	Ser	Leu	Asn	Gln	Ala	Glu	Ala	Asn	Phe	Phe
65				70					75					80	
Ile	Glu	Ala	Lys	Ala	Ala	Ser	Ser	Asn	Trp	Val	Pro	Lys	Ala	His	Ala
				85					90					95	
Asp	Pro	Asp	Thr	Leu	Pro	Trp	Ser	Lys	Glu	Pro	Pro	Arg	Ala	Ala	Thr
			100					105					110		
Ala	Gly	Gln	Gln	Trp	Ala	Leu	Gln	Thr	Val	Leu	Leu	Glu	Val	Leu	Asn
		115					120					125			
Arg	Phe	Met	Pro	Pro	Pro	Asn	Asn	Thr	Pro	Gly	Arg	Thr	Phe	Gly	Arg
	130						135				140				
Thr	Leu	Ser	Gly	Pro	Ser	Gly	Leu	Ser	Arg	Pro	Thr	Ser	Thr	Asn	Thr
145					150					155				160	
Lys	Arg	Lys	Asp	Glu	Pro	Ala	Asn	Val	Thr	Phe	Ala	Asp	Pro	Pro	Lys
			165						170				175		
Arg	Ser	Leu	Thr	Arg	Ser	Ala	Thr	Gly	Pro	Pro	Ile	His	Gly	Ala	Ala
		180						185					190		
Ile	Pro	Leu	Lys	Phe	Pro	Asp	Pro	Val	Asn	Thr	Gly	Ser	Lys	Arg	Pro
	195						200					205			
Ser	Leu	Glu	Ser	Glu	Asn	Leu	Asn	Gln	Cys	Thr	Lys	Arg	Ala	Lys	Gly
	210					215					220				
Lys	Leu	Ser	Asp	Asn	Val	Ala	Ala	Ala	Ala	Ala	Pro	Pro	Val	Pro	Ile
225					230					235				240	

INC. BENTON & BOWLES

Ala Ser Ala Leu Asp Lys Val Pro Thr Arg Arg His Ala Asn Thr Arg
 245 250 255
 Asp Pro Thr Ala Thr Gly His Arg Arg Ala Asp Gln Val Asp Ser Phe
 260 265 270
 Asp Thr Ser Gln Gly Thr Ser Tyr Gly Ser Ser Val Phe Ser Ala Cys
 275 280 285
 Arg His Asn Gln Ser Thr Thr Gln Ser Ser Phe Glu Ala Pro Pro Ser
 290 295 300
 Gln Pro Arg Glu Lys Arg Pro Val Asp Ala Thr Val Phe Glu Ala Gly
 305 310 315 320
 His Leu Ile Glu Ser Pro Ser Lys Gly Arg Thr Thr Lys Ser His Ile
 325 330 335
 Asp Asn Gln Pro Leu Ser Ser Ser Ser Gln Gly Glu Thr Ser Phe Ser
 340 345 350
 Thr Tyr Tyr Glu Ser Phe Pro Ser Ser Gly Gly Glu Gly Ala Ile Pro
 355 360 365
 Glu Pro Ser Arg Ser Asn Gly Leu Ala Arg Ser Glu Glu Ser Ala Arg
 370 375 380
 Ser Gln Val Gln Val His Ala Pro Val Val Ala Ala Arg Leu Arg Asn
 385 390 395 400
 Ile Trp Pro Lys Phe Pro Lys Trp Leu His Glu Ala Pro Leu Ala Val
 405 410 415
 Ala Trp Glu Val Thr Arg Leu Phe Met His Cys Lys Val Asp Leu Glu
 420 425 430
 Asp Glu Ser Leu Gly Leu Lys Tyr Asp Pro Ser Trp Ser Thr Ala Arg
 435 440 445
 Asp Val Thr Asp Ile Trp Lys Thr Leu Tyr Arg Leu Asp Ala Phe Arg
 450 455 460
 Gly Lys Pro Phe Pro Glu Lys Pro Pro Asn Asp Val Phe Val Thr Ala
 465 470 475 480
 Met Thr Gly Asn Phe Glu Ser Lys Gly Ser Ala Val Val Leu Ser Ala
 485 490 495
 Val Leu Asp Tyr Asn Pro Asp Asn Ser Pro Thr Ala Pro Leu Tyr Leu
 500 505 510
 Val Lys Leu Lys Pro Leu Met Phe Glu Gln Gly Cys Arg Leu Thr Arg
 515 520 525
 Arg Phe Gly Pro Asp Arg Phe Phe Glu Ile Leu Ile Pro Ser Pro Thr
 530 535 540

HCG PATENTED & REGISTERED IN THE U.S.A.

Ser Thr Ser Pro Ser Val Pro Pro Val Val Ser Lys Gln Pro Gly Ala
 545 550 555 560
 Val Glu Glu Val Ile Gln Trp Leu Thr Met Gly Gln His Ser Leu Val
 565 570 575
 Gly Arg Gln Trp Arg Ala Phe Phe Ala Lys Asp Ala Gly Tyr Arg Lys
 580 585 590
 Pro Leu Arg Glu Phe Gln Leu Arg Ala Glu Asp Pro Lys Pro Ile Ile
 595 600 605
 Lys Glu Arg Val His Phe Phe Ala Glu Thr Gly Ile Thr Phe Arg Pro
 610 615 620
 Asp Val Phe Lys Thr Arg Ser Val Val Pro Ala Glu Glu Pro Val Glu
 625 630 635 640
 Gln Arg Thr Glu Phe Lys Val Ser Gln Met Leu Asp Trp Leu Leu Gln
 645 650 655
 Leu Asp Asn Asn Thr Trp Gln Pro His Leu Lys Leu Phe Ser Arg Ile
 660 665 670
 Gln Leu Gly Leu Ser Lys Thr Tyr Ala Ile Met Thr Leu Glu Pro His
 675 680 685
 Gln Ile Arg His His Lys Thr Asp Leu Leu Ser Pro Ser Gly Thr Gly
 690 695 700
 Glu Val Met Asn Asp Gly Val Gly Arg Met Ser Arg Ser Val Ala Lys
 705 710 715 720
 Arg Ile Arg Asp Val Leu Gly Leu Gly Asp Val Pro Ser Ala Val Gln
 725 730 735
 Gly Arg Phe Gly Ser Ala Lys Gly Met Trp Val Ile Asp Val Asp Asp
 740 745 750
 Thr Gly Asp Glu Asp Trp Ile Glu Thr Tyr Pro Ser Gln Arg Lys Trp
 755 760 765
 Glu Cys Asp Phe Val Asp Lys His Gln Arg Thr Leu Glu Val Arg Ser
 770 775 780
 Val Ala Ser Glu Leu Lys Ser Ala Gly Leu Asn Leu Gln Leu Leu Pro
 785 790 795 800
 Val Leu Glu Asp Arg Ala Arg Asp Lys Val Lys Met Arg Gln Ala Ile
 805 810 815
 Gly Asp Arg Leu Ile Asn Asp Leu Gln Arg Gln Phe Ser Glu Gln Lys
 820 825 830
 His Ala Leu Asn Arg Pro Val Glu Phe Arg Gln Trp Val Tyr Glu Ser
 835 840 845



Tyr Ser Ser Arg Ala Thr Arg Val Ser His Gly Arg Val Pro Phe Leu
 850 855 860
 Ala Gly Leu Pro Asp Ser Gln Glu Glu Thr Leu Asn Phe Leu Met Asn
 865 870 875 880
 Ser Gly Phe Asp Pro Lys Lys Gln Lys Tyr Leu Gln Asp Ile Ala Trp
 885 890 895
 Asp Leu Gln Lys Arg Lys Cys Asp Thr Leu Lys Ser Lys Leu Asn Ile
 900 905 910
 Arg Val Gly Arg Ser Ala Tyr Ile Tyr Met Ile Ala Asp Phe Trp Gly
 915 920 925
 Val Leu Glu Glu Asn Glu Val His Val Gly Phe Ser Ser Lys Phe Arg
 930 935 940
 Asp Glu Glu Glu Ser Phe Thr Leu Leu Ser Asp Cys Asp Val Leu Val
 945 950 955 960
 Ala Arg Ser Pro Ala His Phe Pro Ser Asp Ile Gln Arg Val Arg Ala
 965 970 975
 Val Phe Lys Pro Glu Leu His Ser Leu Lys Asp Val Ile Ile Phe Ser
 980 985 990
 Thr Lys Gly Asp Val Pro Leu Ala Lys Lys Leu Ser Gly Gly Asp Tyr
 995 1000 1005
 Asp Gly Asp Met Ala Trp Val Cys Trp Asp Pro Glu Ile Val Asp Gly
 1010 1015 1020
 Phe Val Asn Ala Glu Met Pro Leu Glu Pro Asp Leu Ser Arg Tyr Leu
 1025 1030 1035 1040
 Lys Lys Asp Lys Thr Thr Phe Lys Gln Leu Met Ala Ser His Gly Thr
 1045 1050 1055
 Gly Ser Ala Ala Lys Glu Gln Thr Thr Tyr Asp Met Ile Gln Lys Ser
 1060 1065 1070
 Phe His Phe Ala Leu Gln Pro Asn Phe Leu Gly Met Cys Thr Asn Tyr
 1075 1080 1085
 Lys Glu Arg Leu Cys Tyr Ile Asn Asn Ser Val Ser Asn Lys Pro Ala
 1090 1095 1100
 Ile Ile Leu Ser Ser Leu Val Gly Asn Leu Val Asp Gln Ser Lys Gln
 1105 1110 1115 1120
 Gly Ile Val Phe Asn Glu Ala Ser Trp Ala Gln Leu Arg Arg Glu Leu
 1125 1130 1135
 Leu Gly Gly Ala Leu Ser Leu Pro Asp Pro Met Tyr Lys Ser Asp Ser
 1140 1145 1150

Trp Leu Gly Arg Gly Glu Pro Thr His Ile Ile Asp Tyr Leu Lys Phe			
1155	1160	1165	
Ser Ile Ala Arg Pro Ala Ile Asp Lys Glu Leu Glu Ala Phe His Asn			
1170	1175	1180	
Ala Met Lys Ala Ala Lys Asp Thr Glu Asp Gly Ala His Phe Trp Asp			
185	1190	1195	1200
Pro Asp Leu Ala Ser Tyr Tyr Thr Phe Phe Lys Glu Ile Ser Asp Lys			
1205	1210	1215	
Ser Arg Ser Ser Ala Leu Leu Phe Thr Thr Leu Lys Asn Arg Ile Gly			
1220	1225	1230	
Glu Val Glu Lys Glu Tyr Gly Arg Leu Val Lys Asn Lys Glu Met Arg			
1235	1240	1245	
Asp Ser Lys Asp Pro Tyr Pro Val Arg Val Asn Gln Val Tyr Glu Lys			
1250	1255	1260	
Trp Cys Ala Ile Thr Pro Glu Ala Met Asp Lys Ser Gly Ala Asn Tyr			
265	1270	1275	1280
Asp Ser Lys Val Ile Arg Leu Leu Glu Leu Ser Phe Leu Ala Asp Arg			
1285	1290	1295	
Glu Met Asn Thr Trp Ala Leu Leu Arg Ala Ser Thr Ala Phe Lys Leu			
1300	1305	1310	
Tyr Tyr His Lys Ser Pro Lys Phe Val Trp Gln Met Ala Gly Arg Gln			
1315	1320	1325	
Leu Ala Tyr Ile Lys Ala Gln Met Thr Ser Arg Pro Gly Glu Gly Ala			
1330	1335	1340	
Pro Ala Leu Met Thr Ala Phe Met Tyr Ala Gly Leu Met Pro Asp Lys			
345	1350	1355	1360
Lys Phe Thr Lys Gln Tyr Val Ala Arg Leu Glu Gly Asp Gly Ser Glu			
1365	1370	1375	
Tyr Pro Asp Pro Glu Val Tyr Glu Val Leu Gly Asp Asp Asp Phe Asp			
1380	1385	1390	
Gly Ile Gly Phe Thr Gly Asn Gly Asp Tyr			
1395	1400		

BIBLIOGRAFIA

- Baulcombe, D.C. (1996) *Plant Mol. Biol.* 32, 79-88.
- Cogoni, C. et al. (1996) *EMBO J.* 15, 3153-3163.
- Cogoni, C. and Macino, G. (1997) *Proc. Natl. Acad. Sci. U.S.A.* 94: 10223-10238.
- Cabibbo, A. et al. (1991) *Fungal Genetic Newsl.*, 38: 68-70.
- Davis, R.H. and De Serres, F.J. (1970) *Methods Enzymol.* 17: 79-143.
- de Carvalho Niebel, F. et al. (1995), *Plant Cell* 7: 347-358.
- Dehio, C., and Schell, J. (1994). *Proc. Natl. Acad. Sci. U.S.A.* 91: 5538-5542.
- Dorer, D.R. and Henikoff, S. (1994). *Cell* 77, 993-1002.
- Elkind, Y., et al. (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87: 9057-9061.
- Flavell, R.B. (1994) *Proc. Nat. Acad. Sci. U.S.A.*, 91: 3490-3496.
- Garrick D., et al. (1998). *Nature Genetics* 18, 56-59
- Lindbo, J.A. et al. (1993) *Plant Cell* 5:1749-1759.
- Maniatis, S.T. et al. (1982) *Molecular Cloning-A Laboratory Manual*. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.
- Morelli, G. et al. (1993) *Methods Enzymol.* 214:412-424.
- Mittelstein Scheid, O. et al. (1994) *Mol. Gen. Genet.* 244: 325-330.
- Pal-Bhadra, M., et al. (1997). *Cell* 90, 479-490.
- Palauqui, J.C. et al. (1997) *EMBO J.* 16: 4738-4745.
- Pirrotta, V. (1997). *TIG* 13, 314-318.
- Schiebel et al. (1998) *The Plant Cell* 10:2087-2101.
- Schmidhauser, T.J. et al. (1990) *Mol. Cell. Biol.* 10:5064-5070.
- Staben, C. et al. (1989) *Fungal Genetics Newsl.* 36:79-81.
- Stam, M. et al. (1997) *Annals of Botany* 79: 3-12.

Stam, M. et al. (1997), Plant Journal 1: 63-82.

van Blokland, R. et al. (1994), Plant. 6, 861-887.

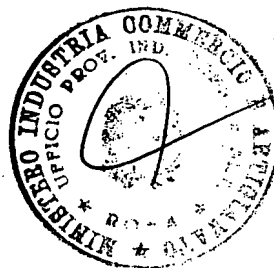
Vaucheret, H. (1993), C.R. Acad. Sci. Paris, Sciences de la vie/Life sciences 316, 1471-1483.

Voellmer, L.J. and Yanofsky, C. (1987) Proc. Natl. Acad. Sci. USA 83, 4869-4873.

Wallrath, L.L. and Elgin, S.C.R. (1995). Genes & Development 9, 1263-1277.

UN MANDATARIO
per se e per gli altri
Olga Capasso
(p.d. 820 B)

Olga Capasso



ING. BARILE & ZANARDI ROMA SpA

RM 99 A 000117

RIVENDICAZIONI

1. Sequenza nucleotidica codificante per una proteina caratterizzata dal fatto di avere un'attività di silenziamento e di comprendere un dominio di RNA polimerasi RNA-dipendente, in cui il dominio è omologo almeno per il 30% con la sequenza amminoacidica dall'amminoacido 710 all'amminoacido 1282 di SEQ ID No.1.
2. Sequenza nucleotidica codificante per una proteina caratterizzata dal fatto di avere un'attività di silenziamento e di comprendere un dominio di RNA polimerasi RNA-dipendente secondo la rivendicazione 1 in cui il dominio è omologo almeno per il 40% con la sequenza amminoacidica dall'amminoacido 710 all'amminoacido 1282 di SEQ ID No.1.
3. Sequenza nucleotidica codificante per una proteina caratterizzata dal fatto di avere un'attività di silenziamento e di comprendere un dominio di RNA polimerasi RNA-dipendente secondo la rivendicazione 2 in cui il dominio è omologo almeno per il 50% con la sequenza amminoacidica dall'amminoacido 710 all'amminoacido 1282 di SEQ ID No.1.
4. Sequenza nucleotidica codificante per una proteina caratterizzata dal fatto di avere un'attività di silenziamento e di comprendere un dominio di RNA polimerasi RNA-dipendente secondo la rivendicazione 3 in cui il dominio è la sequenza amminoacidica dall'amminoacido 710 all'amminoacido 1282 di SEQ ID No.1.

ING. BAZZANO & ZUCCHETTI SPA

5. Sequenza nucleotidica codificante per una proteina caratterizzata dal fatto di avere un'attività di silenziamento e di comprendere un dominio di RNA polimerasi RNA-dipendente secondo la rivendicazione 4 in cui detta sequenza nucleotidica codifica per una proteina avente la sequenza amminoacidica di SEQ ID No.1, o porzioni funzionali di essa.
6. Sequenza nucleotidica codificante per una proteina caratterizzata dal fatto di avere un'attività di silenziamento e di comprendere un dominio di RNA polimerasi RNA-dipendente secondo la rivendicazione 5 in cui detta sequenza nucleotidica è la sequenza nucleotidica di SEQ ID No.1 o la sua sequenza complementare.
7. Vettore di espressione comprendente sotto il controllo di un promotore esprimibile in batteri la sequenza nucleotidica secondo una qualsiasi delle rivendicazioni da 1 a 6.
8. Vettore di espressione comprendente sotto il controllo di un promotore esprimibile in piante o in parti di esse la sequenza nucleotidica secondo una qualsiasi delle rivendicazioni da 1 a 6, in orientamento senso o antisenso.
9. Vettore di espressione comprendente sotto il controllo di un promotore esprimibile in funghi la sequenza nucleotidica secondo una qualsiasi delle rivendicazioni da 1 a 6, in orientamento senso o antisenso.
10. Vettore di espressione comprendente sotto il controllo di un promotore esprimibile in animali la sequenza nucleotidica secondo una

qualsiasi delle rivendicazioni da 1 a 6, in orientamento senso o antisenso.

11. Organismo procariotico trasformato con il vettore di espressione in batteri secondo la rivendicazione 7.

12. Pianta o parti di essa trasformata con il vettore di espressione in piante secondo la rivendicazione 8.

13. Pianta mutata nella sequenza nucleotidica secondo una qualsiasi delle rivendicazioni da 1 a 6 con un'attività di silenziamento ridotta o inibita.

14. Fungo trasformato con il vettore di espressione in funghi secondo la rivendicazione 9.

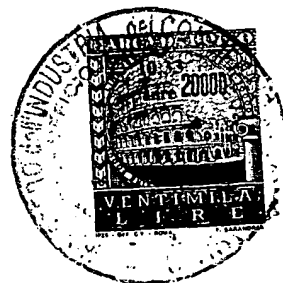
15. Fungo mutato nella sequenza nucleotidica secondo una qualsiasi delle rivendicazioni da 1 a 6 con un'attività di silenziamento ridotta o inibita.

16. Animale non umano trasformato con il vettore di espressione in animali secondo la rivendicazione 10.

17. Animale non umano mutato nella sequenza nucleotidica secondo una qualsiasi delle rivendicazioni da 1 a 6 con un'attività di silenziamento ridotta o inibita.

18. Proteina caratterizzata dal fatto di avere un'attività di silenziamento e di comprendere un dominio di RNA polimerasi RNA-dipendente, in cui il dominio è omologo almeno per il 30% con la sequenza amminoacidica dall'amminoacido 710 all'amminoacido 1282 di SEQ ID No.1.

19. Proteina caratterizzata dal fatto di avere un'attività di silenziamento e di comprendere un dominio di RNA polimerasi RNA-dipendente secondo la rivendicazione 18 in cui il dominio è omologo almeno per il 40% con la sequenza amminoacidica dall'amminoacido 710 all'amminoacido 1282 di SEQ ID No.1.



20. Proteina caratterizzata dal fatto di avere un'attività di silenziamento e di comprendere un dominio di RNA polimerasi RNA-dipendente secondo la rivendicazione 19 in cui il dominio è omologo almeno per il 50% con la sequenza amminoacidica dall'amminoacido 710 all'amminoacido 1282 di SEQ ID No.1.

21. Proteina caratterizzata dal fatto di avere un'attività di silenziamento e di comprendere un dominio di RNA polimerasi RNA-dipendente secondo la rivendicazione 20 in cui il dominio è la sequenza amminoacidica dall'amminoacido 710 all'amminoacido 1282 di SEQ ID No.1.

22. Proteina caratterizzata dal fatto di avere un'attività di silenziamento e di comprendere un dominio di RNA polimerasi RNA-dipendente secondo la rivendicazione 21 comprendente la sequenza amminoacidica di SEQ ID No.1, o porzioni funzionali di essa.

23. Uso della sequenza nucleotidica secondo una qualsiasi delle rivendicazioni da 1 a 6 per la modulazione del silenziamento genico in piante, animali e funghi.

Roma, 22 FEB. 1999

p.p.: Università degli Studi di Roma "La Sapienza"

Ing. Barzanò & Zanardo Roma S.p.A.

dpk capasso

UN MANDATARIO
per se e per gli altri
Olga Capasso
(N° d'iscr. 820 B)



ING. BARZANÒ & ZANARDO ROMA SpA

1/5

RM 99 A 000117

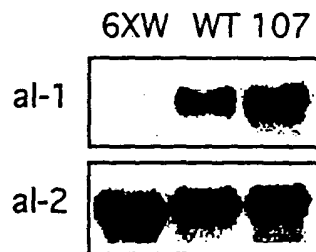
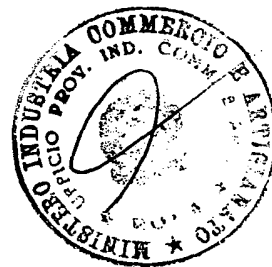


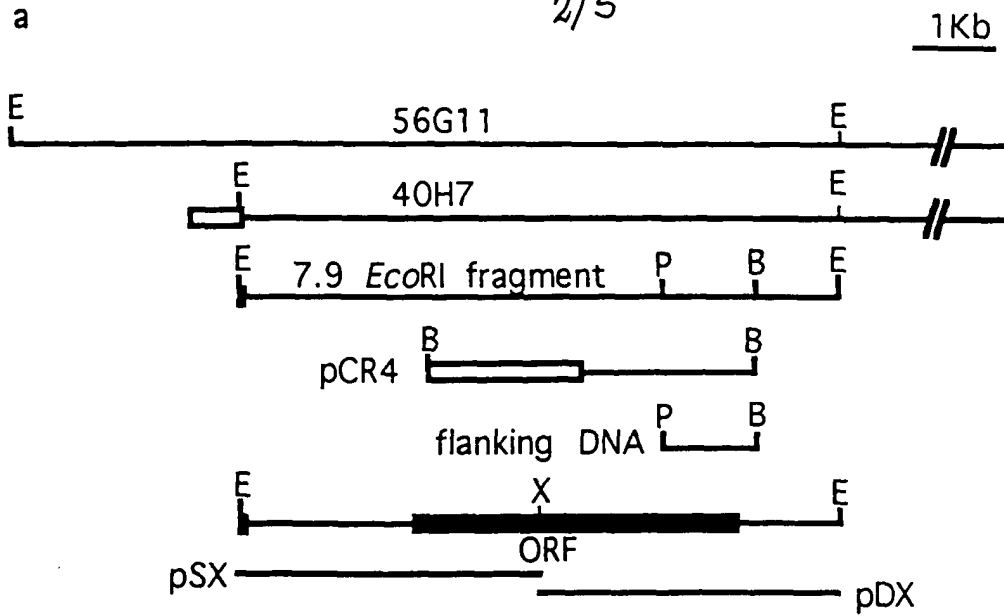
FIG. 1



p.p.: UNIVERSITA' DEGLI STUDI DI ROMA "LA SAPIENZA"
ING. BARZANO' & ZANARDO ROMA S.p.A.

olga capasso

UN MANDATARIO
per se e per gli altri
Olga Capasso
(N° d'iscr. 820 B)



b

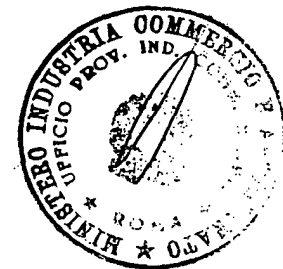
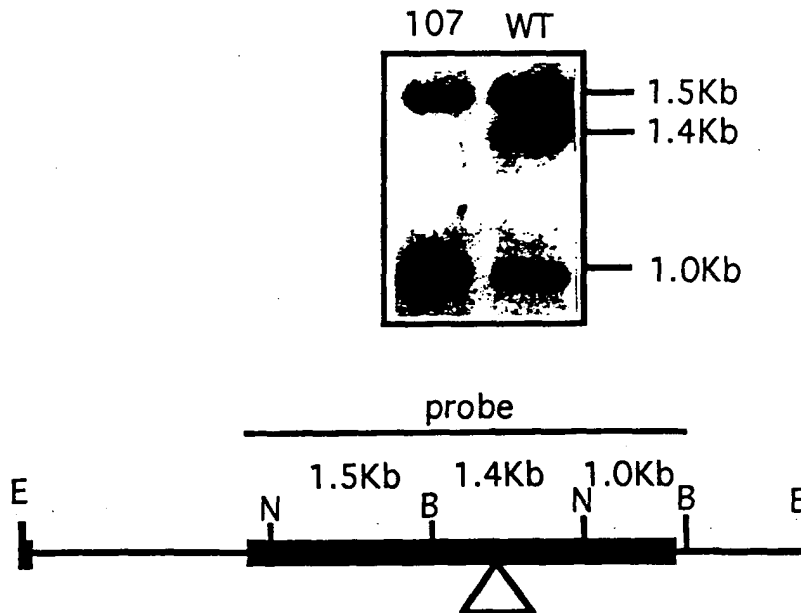


FIG. 2

p.p.: UNIVERSITA' DEGLI STUDI DI ROMA "LA SAPIENZA"
ING. BARZANO' & ZANARDO ROMA S.p.A.

Olga Capasso

UN MANDATARIO
per se e per gli altri
Olga Capasso
(N° d'iscr. 820 B)

3/5

RM 99A000117

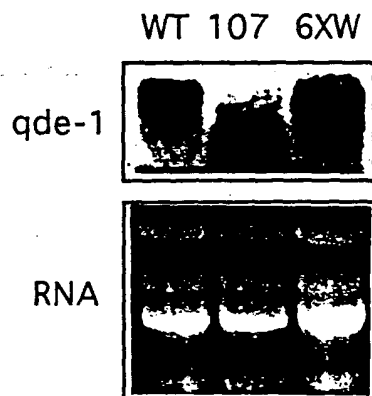


FIG. 3



p.p.: UNIVERSITA' DEGLI STUDI DI ROMA "LA SAPIENZA"
ING. BARZANO' & ZANARDO ROMA S.p.A.

olga capasso

UN MANDATARIO
per se e per gli altri
Olga Capasso
(N° d'iscr. 820 B)

4/5

R M 99 A 000117

MNPITPRKRN SPVEEIIINRL NNDYNLGLQC VADTTLTTPHR PKELAESDED FGRHDKIYRA	60
LNFLYWRKDD SLNQAEANFF IEAKAASSNW VPKAHADPDT LPWSKEPPRA ATAGQQWALQ	120
TVLLEVLNRF MPPPNNTPGR TFGRTLSGPS GLSRPTSTNT KRXDEPANVT FADPPKRSLT	180
RSATGPPIHG AAIPLKFPDP VNTGSKRPSL ESENLNQCTK RAKGKLSDNV AAAAAPPVPI	240
ASALDKVPTR RHANTRDPTA TGHRRADQVD SFDTSQGTSY GSSVFSACRH NQSTTQSSFE	300
APPSQPREKR PVDATVFEAG HLIESPSKGR TTKSHIDNQP LSSSSQGETS FSTYYESFPS	360
SGGEGAIPEP SRSNGLARSE ESARSQVQVH APVVAARLRN IWPKFPKWLH EAPLAVAWEV	420
TRLFMHCKVD LEDESLGLKY DPSWSTARDV TDIWKTLYRL DAFRGKPFPE KPPNDVFVTA	480
MTGNFESKGS AVVLSAVLDY NPDNSPTAPL YLVKLKPLMF EQGCRLTRRF GPDRFFEILI	540
PSPTSTSPSV PPVVSQPGA VEEVIQWLTM GQHSVLGRQW RAFFAKDAGY RKPLREFQLR	600
AEDPKPIIKE RVHFFAETGI TFRPDVFKTR SVVPAEEPVE QRTEFKVSQM LDWLLQLDNN	660
TWQPHLKLFS RIQLGLSKTY AIMTLEPHQI RHHKTDLLSP SGTGEVMNDG VGPMRSVAK	720
<u>RIRDVLGLGD VPSAVOGRFG SAKGMWVIDV DDTGDEDWIE TYP SORKWEC DEVDKHORTL</u>	780
<u>EVRSVASELK SAGLNLOLLP VLEDRARDKV KMROAIGDRL INDLORQFSE OKHALNRPVE</u>	840
<u>FRQWVYESYS SRATRVSHGR VPFLAGLPDS OEETLNELMN SGFDPKKOKY LODIAWDLOK</u>	900
<u>RKCDTLKSKL NIEVGRSAVI YMIADFVGVL EENEVHVGF SFRDEEESF TLISDCDVLV</u>	960
<u>ARSPAHEPSD IORVRAVEKP ELHSLKDVII FSTKGDVPLA KKLGGDYDG DMAWVCWDPE</u>	1020
<u>IVDGFVNAEM PLEPDLRYL KKDKTTEKOL MASHGTGSAA KEOTTYDMIO KSFHFALOPN</u>	1080
<u>ELGMCTNYKE RLCYINNSVS NKPAILSSL VGNLVDOSKO GIVFNEASWA OLRRELLGGA</u>	1140
<u>LSLPDPMYKS DSWLGRGEPT HIIDYLKFSI ARPAIDKELE AFHNAMKAAK DTEDGAHFW</u>	1200
<u>PDLASYTTFE KEISDKSRSS ALLETTLKNR IGEVEKEYGR LVKNKEMRDS KDPYPVRVNO</u>	1260
<u>VYEKWCAITP EAMDKSGANY DSKVIRLLEL SFLADREMNT WALLRASTAF KLYYHKSPKF</u>	1320
VWQMAGRQLA YIKAQMTSRP GEGAPALMTA FMYAGLMPDK KFTKQYVARL EGDGSEYPDP	1380
EVYEVLGDDD FDGIGFTGNG DY	

FIG. 4



p.p.: UNIVERSITA' DEGLI STUDI DI ROMA "LA SAPIENZA"
ING. BARZANO' & ZANARDO ROMA S.p.A.

Olga Capasso

UN MANDATARIO
per se e per gli altri
Olga Capasso
(N° d'iscr. 820 B)



araB GTGMSISLAFKQVAQKCG--ISHVPSAFCEFYGGYNG--VIAVLRSSFRKLSL-----FD 606
 RdRP GIGKISGDFHFAVASKCG--LOYTPSAFCIFYGGYNG--VVGVDSDSSMKLSL-----FK 577
 pom GVMASLSVIRELSLEVKNHDMFSAFCFFMGYNG--VLSLAPPTKLEYHQGNLVFPFR 672
 elegl GCGRISIKLATHISKILO--LKEVPACEGVFFKGFNG--LIVIDEITIDDIINMP-KVIFPK 923
 qdel GVGRMSRSVAKFIRDVLG--LGDVPSAVCGFFGSANGMWVIDVIDTGDDEWIET-----YP 763

araB SMLKFDSN-----NRMLNVTWRT-ESMECFLLHFEILCLLSTLCIEDAMFEA-----MQAV 655
 RdRP SMSKYESD-----NIKLDVLGWS-KYQPCYLHFLQITLLSTLCGVKDEVLEQ-----KQKE 626
 pom SCQDKFKSF-----HSTLEVIKIS-RFSNAHLNLMOLITLLEGLGVKTVFLE-----LTRS 721
 elegl SQCKFEGEGGELQDEYLEVVKYA-MPSEVCLHPPFITLLDQVSEKQSASSHRR--ITNRV 980
 qdel SQRFWECDVFDKHQRTLEVRSVASELKSAGLNLLOLPLVEDRARDKVKMQAIGDRLIND 823

araB HLSMLGNMLEDRDAALNVCKLS-GENSKMLLVKMMLLQ----CYAPSSSEYYSMLLRVHH 710
 RdRP AVDQLDAILHDSLKAQEAELMSPGENT-NILKAMLMC-----GYKPDABEFLSMMLQTFR 681
 pom QLSKMNESINSKQKSILMRDNVDEYHSTLIADFIQA----GFLERDDAFTENILNLYY 777
 elegl HYYLERELCISLNMNLINENAAEELVIRTLAIDWNAASKRAGFELSVDLDRDLFSIY 1040
 qdel LQRQFSEQKHALNRPVEFRQWVYESYSSRATRVSNGRVPFLAGLPDSQETINFTNSGF 883

araB ESQSELKSRCRILVPI--GRILICCMDEMCILE--YGQVYVRVTLTKAELKSRDQSYFR 766
 RdRP ASKLLDLRTSRIFTFN--GRTMGCLDESTLE--YGQVYVQFTGAGHGEFSDDLHPFN 737
 pom EWVLRRIKERQKVSVP--GAYLLGVADETCILKGHDDAVLSVPEIFIQITDTSTSGS 835
 elegl RYNIHHISAKIFLEPSLGSFMYGVVDETCLLQ--YGQVFIQYSPSIRQTSNRP----- 1093
 qdel DPKKQKYLODIAWDLOR-KCDTLKSKLNIRVGRSAIYIADFWGVLEENEVHVGFSSK 942

araB KIDEETS--VV-IGKVVTNPNCLHPGDIRVLDIYEVHFEEKGYLDCIIFPQGERPHF 823
 RdRP NSRSTNSNFIL-KGNVVAKNPCLHPGDIRVLKAVNVRAH--MVLCVVFQPKGKRPHE 794
 pom YSTGKLKTRVI-VGLCLVARHSLHPGVVRCKAVRCDEIMH--LKNVIVFTTCDPSIF 892
 elegl --ILKT-----GKVLITRHPCHVPGDVRFDAVWQPALAH--LVQVVFPCPCPHE 1142
 qdel FRDEEESFTLSDCDVIVARSAHFPSDIQVRRAVFKPELHS--LKDVITESTGDEVELA 1000

araB NECSGGDLGDGQFFVSWLEKLI--SEMDPPDYAGSRPRLMCHD----- 866
 RdRP NECSGDLGDGIFVCDQDMIFPRQVQPEYPPAPSIQLDHE----- 837
 pom AMCSGGDLGDGENTVIWDQRILIKIVNYEPLLESSPKKSIFLEG----- 937
 elegl DEMAGSGLDGCESTINDQEMLDYNEEA VFPSSSAAEEDKE----- 1185
 qdel KKLGGGLYDGDMAWVCDPEIVDGFVNAEPLPDLRYLKKKTTFKQLMASHGTGSAA 1060

araB ---VLEEIHKFEVDNMISDTLGVIST--AHIVHA-----DRDPEKARSQKOLELANLHSR 917
 RdRP ---VHIEVEEYETNYIVNDSLGHIAH--AHVVEA-----DREPDAMSDPCKKLAELESI 888
 pom ---KPLILSVKEFFVNYIKYDSLGLISH--AWKAWH--DHENNEEGIFGNVCELEAEMHSK 992
 elegl ---PITUDMVEFFLRYLQODSIGRMH--AHLAYE-----DLHG-LFHENCHAL-LKCAV 1234
 qdel KEQTTYDMIQSSEHFAIQPNFLGMCNHYKERLCYINNSVSNKEAILLSLVGNLVDQSKQ 1120

araB AVDEANTCAPAEMPYAKF--REFPDELERFERPTTISESVFGNLYFAVKSSLAQRKPEE 976
 RdRP AVDEPKTGVPAPPSQLR--KEYPDEMDPDITSISEFVIGHLEPHVKDKAPQASSIAI 947
 pom AVDEAFSGWACKMQAKYH--KRYPDFOHTKTRSFSETAVGNIFFYAARFQRESGRPAI 1051
 elegl AVDEPESGVFAEPLSSFEQCEMTPIYMSGGPMYSTEELNQLHFAARKVEEVLEEFEL 1294
 qdel GIVNEASWAQLRRELGGALSPLPPIYSDSWLGRCEPHTIIDYLFSTIARPAIDKELE 1180

araB SEDTVA---YDVTLEEAGEES-F-----IETAKAHDRMGEKLTSLIYFGAANEEIILT 1027
 RdRP FTRDVARRSYDADMEVDGED-Y-----IDEAFDYKTEYDNKLGNDYGTIKTEAFILS 1001
 pom YNPIMN-TVYUPCMKLPREKTEY-----LNVAEEVKKHNDNERSIARFDISTEYEVYT 1105
 elegl RGSVFER-EYKCLICPEDVDVFFGNEIKLVQTLTLRDETVDRYQQLDEYGIIDEASVVS 1353
 qdel AFHNAMKAAKETEDGAHFWDPLAS--YTFEKEISDKSRSSALLFTTLKNRIGEVKEY 1238

araB G-IDTKEMYLA-RNRRNGDMKDRITLSVKDLHKEAMGWEEKSCDE----- 1073
 RdRP GGIIMASKTF---ERRKD---AEALSVAVRALRKEARAWKRRNDID----- 1042
 pom AFILFKDDLAK---TVNEIG-LREEVSFQFDLLIKYTOEYLEKCAL----- 1149
 elegl GHAASIKRLAGMERDYSFYHTDKVVELRYEKLYAVFRAKEFEFFGEEINIENDGKNTR 1413
 qdel GREYENKEMRD---SKDPPVRVNQVYEKWCATPEAMDKSGANYDSK----- 1283

FIG. 5

